

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____
Date: _____ Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 01-10-03
Searcher: Beverly E. 4494
Terminal time: 20
Elapsed time: _____
CPU time: _____
Total time: 23
Number of Searches: _____
Number of Databases: 1

Search Site

____ STIC
____ CM-1
____ Pre-S

Type of Search

____ N.A. Sequence
____ A.A. Sequence
____ Structure
____ Bibliographic

Vendors

____ IG
____ STN
____ Dialog
____ APS
____ Geninfo
____ SDC
____ DARC/Questel
☒ Other CGN

us-09-873-737a-5.rge

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

(without alignments)
17074.171 Million cell updates/sec

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 Bqgs, 4109280

Maximum DB seq length: 20000000000

Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_cm:*
21: em_ov:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_hgo_hum:*
40: em_hgo_mus:*
41: em_hgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID
---------------	-------	----------------	--------	----	----

Description
AF104260 Homo sapi

1	3359	96.7	3362	9	AF104260
2	3344.6	96.1	3399	9	BC028581
3	3337.6	96.3	3421	9	AF387507
4	2513.6	72.4	2579	9	AK093313
5	2211	66.6	2328	9	AF264004
6	1839.2	53.0	3910	10	AB032604
7	1839.2	53.0	4064	10	AF438405
8	1067.7	30.7	3273	5	AF33362
9	830.8	23.9	193697	2	AC025837
10	830.8	23.9	198104	2	AC127071
11	679.2	19.6	3306	9	BC031060
12	679.2	19.6	3138	9	AK093475
13	613.2	17.7	3456	9	AK096837
14	565	16.3	3057	3	AX014899
15	550.4	15.9	3580	6	AX381696
16	495.8	14.3	2605	3	AX014900
17	459.4	13.2	501	6	AX395819
18	458.4	13.2	4913	10	AX098665
19	456.6	13.2	3610	9	BC025935
20	438.8	12.5	3620	9	AK056418
21	434.2	10.9	1723	3	AY014901
22	378.4	10.9	3649	6	AX247684
23	377	10.9	3659	10	AF285586
24	377	10.9	2272	9	AK027497
25	329	9.5	2761	3	AF145680
26	310.2	8.9	3047	3	AK001213
27	253.4	7.3	1733	3	AC090965
28	238.8	6.9	61995	2	AC090965
29	203.4	5.9	61995	2	DDNDNRP9
30	186.2	5.4	7781	3	AC005116
31	184.6	5.3	37980	2	AC019704
32	184.6	5.3	78507	2	AC092226
33	184.6	5.3	169210	3	AC092226
34	184.6	5.3	176306	3	AE007186
35	184.6	5.3	276556	3	AE003630
36	183	5.3	901	3	AF344408
37	170.8	4.9	548	6	AC387649
38	170.6	4.9	236516	2	AC111089
39	160.4	4.6	3677	2	AC017885
40	160.4	4.6	3855	2	AB003107
41	141.2	4.1	9359	3	AF104355
42	138.4	4.0	33965	2	AC100141
43	129.6	3.7	66611	2	AC050147
44	126.6	3.6	66611	2	AC050147
45	126.6	3.6	1263	9	AK000397
					BC028581 Homo sapi
					AF387507 Homo sapi
					AK093313 Homo sapi
					AF264004 Homo sapi
					AB032604 Mus muscu
					AF438405 Mus muscu
					AF33362 Danio rer
					AC025837 Homo sapi
					AC127071 Homo sapi
					BC031060 Homo sapi
					AK093475 Homo sapi
					AK096837 Homo sapi
					AY014899 Strongylo
					AX381696 Sequence
					AY014900 Strongylo
					AX385819 Sequence
					AX209866 Sequence
					AB032605 Mus muscu
					BC025935 Homo sapi
					AK056418 Homo sapi
					AY014901 Strongylo
					AX247684 Sequence
					AF285586 Mus muscu
					AK027497 Homo sapi
					AF145680 Drosophill
					AF104354 Drosophill
					AK001213 Homo sapi
					AC090965 Homo sapi
					X94613 D.melanogast
					AC005116 Drosophill
					AC019704 Drosophill
					AC092226 Drosophill
					AC097186 Drosophill
					AE003630 Drosophill
					AF344408 Drosophill
					AX387649 Sequence
					AC111089 Mus muscu
					AC017885 Dros

ALIGNMENTS

RESULT 1		PRI 09-JAN-2002
AF104260	Homo sapiens HIW1 mRNA, complete cds.	linear
LOCUS		
DEFINITION		
ACCESSION		
AF104260.2	GI:18098557	

ORGANISM	REFERENCE AUTHORS TITLE
Homo sapiens	Cox, D. N., Chao, A., Baker, J., Chang, L., Qiao, D. and Lin, H. 1 (Pages 1 to 3362) A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	

Db 1321 GACGACCTATGATTAATTCATTAATAAAGATATAGTTCAAAAGGAGCTTGAGACTGG 1380
 Oy 1491 GTTTAGCTTTGATTTCCACCTACTGTCTCTCAGAGAGAAATTTTGCAACAGAAAAGA 1550
 Db 1381 GTTTAGCTTTGATTTCCACCTACTGTCTCTCAGAGAGAAATTTTGCAACAGAAAAGA 1440
 Oy 1551 TTCCACCAAGTGGAAAAACATTTGATTAACATCCACAAATTTGACAGTTGTGCCAAAAGAA 1610
 Db 1441 TTCCACCAAGTGGAAAAACATTTGATTAACATCCACAAATTTGACAGTTGTGCCAAAAGAA 1500
 Oy 1611 CAAGAGGTGACACATTAATTAAGTTTAAAGCACTAGATTAATCTGGCTGTGATTAAGC 1670
 Db 1501 CAAGAGGTGACACATTAATTAAGTTTAAAGCACTAGATTAATCTGGCTGTGATTAAGC 1560
 Oy 1671 GAGAAATTTTAAAGGAGCCCAATTCATTTGATTAACAAATCTATTTTAAAGTTACACAGCCA 1730
 Db 1561 GAGAAATTTTAAAGGAGCCCAATTCATTTGATTAACAAATCTATTTTAAAGTTACACAGCCA 1620
 Oy 1731 TGGGATGCAAAATGAGAAAAGCAATTAATGATTAAGTGGATGACAGAACTGAAAGCCTAAT 1790
 Db 1621 TGGGATGCAAAATGAGAAAAGCAATTAATGATTAAGTGGATGACAGAACTGAAAGCCTAAT 1680
 Oy 1791 TTAAGGTCTTACGAAAAGGTCAACAGACACCCAGATATGTGTCTGTCTGTGTCAA 1850
 Db 1681 TTAAGGTCTTACGAAAAGGTCAACAGACACCCAGATATGTGTCTGTCTGTGTCAA 1740
 Oy 1851 GTATATGGAAGACAAATTAAGATGCTATTTAAAAATACCTGTGTACATATGCCCCTAACC 1910
 Db 1741 GTATATGGAAGACAAATTAAGATGCTATTTAAAAATACCTGTGTACATATGCCCCTAACC 1800
 Oy 1911 CAAGTCAGT 1970
 Db 1801 CAATGCTGT 1860
 Oy 1971 CAAAGATTTGCCCTACAGATGAACCTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
 Db 1861 CAAAGATTTGCCCTACAGATGAACCTGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2090
 Oy 2031 CCTTGAACCTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
 Db 1921 CCTTGAACCTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2150
 Oy 2091 GGTCAATGCAAGATTTTGTGTCACATCAATGAAGAGATGACCCGCTGTGTCTCAGCT 2040
 Db 1981 GGTCAATGCAAGATTTTGTGTCACATCAATGAAGAGATGACCCGCTGTGTCTCAGCT 2210
 Oy 2151 GCATATTTCAAGATTAAGAGACAGAGAGCTGTATGAGTGGCTCAAAGTCTGCTCAGAGCGG 2210
 Db 2041 GCATATTTCAAGATTAAGAGACAGAGAGCTGTATGAGTGGCTCAAAGTCTGCTCAGAGCGG 2270
 Oy 2211 CTCTGAGAGCTTGAATTAAGTGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
 Db 2101 CTCTGAGAGCTTGAATTAAGTGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2230
 Oy 2271 ATGAGGTGAG 2220
 Db 2161 ATGAGGTGAG 2390
 Oy 2331 ATTTGCTAAAATTCATTTGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 2280
 Db 2221 ATTTGCTAAAATTCATTTGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 2450
 Oy 2391 AAAGAGTGAACACAGATTTTGTCTCAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
 Db 2281 AAAGAGTGAACACAGATTTTGTCTCAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2510
 Oy 2451 GAAACAGTATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 2400
 Db 2341 GAAACAGTATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 2570
 Oy 2511 AGGCTGAGAGAGTGAAGTGTCTCCACACATTAACATGATCACTATTAAGCAACAGCG 2570

[illegible]

RESULT 2					
BC028581	3399 bp	mRNA	linear	PRI	25-APR-2002
LOCUS	BC028581	Homo sapiens, piwi-like 1 (Drosophila), clone MGC:26748			
DEFINITION	IMAGE:4827445, mRNA, complete cds.				
ACCESSION	BC028581				

bb
15 GGGCTGAGGTGCAAGGACGAGGACTAGGCGAGGCGGTCCAGAAATAGAAAACAA 167
|||||
|||||

QY 1248 TTTTGTGAGCCAGCCCAAGAAAGGGGGGCGCTGGGGGGGACATGCGAGGGCTTGCCA 1307
 DB 1155 TTTTGTGAGCCAGCCCAAGAAAGGGGGGCGCTGGGGGGGACATGCGAGGGCTTGCCA 1214
 QY 1308 TGTCTATTCCTGAGCTGTGCTATCTTACAGAGCTTACCTGATTAATAATGCTTAATTTTA 1367
 DB 1215 TGTCTATTCCTGAGCTGTGCTATCTTACAGAGCTTACCTGATTAATAATGCTTAATTTTA 1274
 QY 1368 AGCTGATGAAGAATTAGCCGCTTACCAAGACTACCGAGCAAGGAGCGGTGAAG 1427
 DB 1275 AGCTGATGAAGAATTAGCCGCTTACCAAGACTACCGAGCAAGGAGCGGTGAAG 1334
 QY 1428 TGGGAGAGCTCATTTGATATCATTTCAATAAAGCATTAATGTTCAAGGAGCTTGAGACT 1487
 DB 1335 TGGGAGAGCTCATTTGATATCATTTCAATAAAGCATTAATGTTCAAGGAGCTTGAGACT 1394
 QY 1488 GGGGTTTGAAGCTTGTATTCOAATCTGTCTCTGCGAGGAAGATTTTGCAACAGAAA 1547
 DB 1395 GGGGTTTGAAGCTTGTATTCOAATCTGTCTCTGCGAGGAAGATTTTGCAACAGAAA 1454
 QY 1548 AGATTCAACCAAGGTGGAAAAACATTTGATTACATCCAAATTTGCAAGATTGTCGAAAG 1607
 DB 1455 AGATTCAACCAAGGTGGAAAAACATTTGATTACATCCAAATTTGCAAGATTGTCGAAAG 1514
 QY 1608 AAACAAGAGTGCACCAATTAATTAAGTTTAAGCTACTATTAATCTGGCTGTGATCTATA 1667
 DB 1515 AAACAAGAGTGCACCAATTAATTAAGTTTAAGCTACTATTAATCTGGCTGTGATCTATA 1574
 QY 1668 CGGCAAGAAATTTATGAGAGCGCAATTCATGATACAAATCTATTTAAAGTTAACGCG 1727
 DB 1575 CGGCAAGAAATTTATGAGAGCGCAATTCATGATACAAATCTATTTAAAGTTAACGCG 1634
 QY 1728 CCATGGGAGTCAATATGAGAAAGCAATTAATTAAGTTTAAGCTACTATTAATCTGGCTGTG 1787
 DB 1635 CCATGGGAGTCAATATGAGAAAGCAATTAATTAAGTTTAAGCTACTATTAATCTGGCTGTG 1694
 QY 1788 ACTTAAGAGTCTTACAGCAAAAGGTGACAGAGACCCAGATAGTTGTCTGTCTGTGTGT 1847
 DB 1695 ACTTAAGAGTCTTACAGCAAAAGGTGACAGAGACCCAGATAGTTGTGTCTGTGTGTGT 1754
 QY 1848 CAAGTATTCGAGAGAGCAAAATACAGATGCTATTAATAAATACTGTGTGACAGATTGCCCTA 1907
 DB 1755 CAAGTATTCGAGAGAGCAAAATACAGATGCTATTAATAAATACTGTGTGACAGATTGCCCTA 1814
 QY 1908 CCCCAAGT 1967
 DB 1815 CCCCAAGT 1874
 QY 1968 CTACAAAGATTGCCCTTACAGATGAACTGCAAGATGGGAGAGCTGTGGAGGTGAGCA 2027
 DB 1875 CTACAAAGATTGCCCTTACAGATGAACTGCAAGATGGGAGAGCTGTGGAGGTGAGCA 1934
 QY 2028 TCCCCCTGAGGCTGTGATGATGTTGGCAGTGTGATGTTTACAGATGACATGACGTGGG 2087
 DB 1935 TCCCCCTGAGGCTGTGATGATGTTGGCAGTGTGATGTTTACAGATGACATGACGTGGG 1994
 QY 2088 GGAAGTCAATGCGAGGATTTGTTGCGACAGATCAATGAAGGAGTGAACCGCTGTGTTTCA 2147
 DB 1995 GGAAGTCAATGCGAGGATTTGTTGCGACAGATCAATGAAGGAGTGAACCGCTGTGTTTCA 2054
 QY 2148 GCTGCAATTTTACAGATGAGAGCAGAGACTGTGATGATGAGGCTCAAAATCTGGCTGCAAG 2207
 DB 2055 GCTGCAATTTTACAGATGAGAGCAGAGACTGTGATGATGAGGCTCAAAATCTGGCTGCAAG 2114
 QY 2208 CGGCTCTGAGGAGGTTGGAATAGCTGCAATGAGTACATGCGCAGCGGATCAATCGTGTACC 2267
 DB 2115 CGGCTCTGAGGAGGTTGGAATAGCTGCAATGAGTACATGCGCAGCGGATCAATCGTGTACC 2174
 QY 2268 GCGATGCGATGAGAGAGCGGCGAGCTGAAAAACATGTTGAATCTACAGAGTGGCCAGGTTT 2327
 DB 2175 GCGATGCGATGAGAGAGCGGCGAGCTGAAAAACATGTTGAATCTACAGAGTGGCCAGGTTT 2234
 QY 2328 TGAATGTCTTAAATTCATTTGTGTAGAGTTTAAACACCTGAGATTAACGATTAATTTGTGTGA 2387

DB 2235 TGAATGTCTTAAATTCATTTGTGTAGAGTTTAAACACCTGAGATTAACGATTAATTTGTGTGA 2294
 QY 2388 AGAAAAAGTGAACACCAAGATTTTGTGCTGAGCTGAGAGAGACTTCAGAAATCCACTTC 2447
 DB 2295 AGAAAAAGTGAACACCAAGATTTTGTGCTGAGCTGAGAGAGACTTCAGAAATCCACTTC 2354
 QY 2448 CTGGAACGATTAATGATGATGAGGTTTACAGAGAGATGATGATGATGATGATGATGATGATG 2507
 DB 2355 CTGGAACGATTAATGATGATGAGGTTTACAGAGAGATGATGATGATGATGATGATGATGATG 2414
 QY 2508 GCGAGCTGTGAGAGAGTGTGATGTTTCTCCACACATTAATGATGATGATGATGATGATGATG 2567
 DB 2415 GCGAGCTGTGAGAGAGTGTGATGTTTCTCCACACATTAATGATGATGATGATGATGATGATG 2474
 QY 2568 GGGGCTGAG 2627
 DB 2475 GGGGCTGAG 2534
 QY 2628 ACTGGCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2687
 DB 2535 ACTGGCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2594
 QY 2688 TTTGTTGGCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2747
 DB 2595 TTTGTTGGCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2654
 QY 2748 TCTAACCTGAG 2807
 DB 2655 TCTAACCTGAG 2714
 QY 2808 AGCTTTTATTTATCTTTTATTTTATCTTTTATCTTTTATCTTTTATCTTTTATCTTTTATCT 2867
 DB 2715 AGCTTTTATTTATCTTTTATTTTATCTTTTATCTTTTATCTTTTATCTTTTATCTTTTATCT 2774
 QY 2868 TAGAGATCTAGAGATTTTATCTTTTATCTTTTATCTTTTATCTTTTATCTTTTATCTTTTATCT 2927
 DB 2775 TAGAGATCTAGAGATTTTATCTTTTATCTTTTATCTTTTATCTTTTATCTTTTATCTTTTATCT 2834
 QY 2928 AAAAATTAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 2987
 DB 2835 AAAAATTAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 2894
 QY 2988 TTTGTTTATTTTGAAGAAATGAGATTAAGATCTTGTGATTAATAACAGACTCTGTGAGA 3047
 DB 2895 TTTGTTTATTTTGAAGAAATGAGATTAAGATCTTGTGATTAATAACAGACTCTGTGAGA 2954
 QY 3048 GTATTTGAATGT 3107
 DB 2955 GTATTTGAATGT 3014
 QY 3108 TTTATTAACCTTATTTATCTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 3167
 DB 3015 TTTATTAACCTTATTTATCTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 3074
 QY 3168 GGGCTGAAGTGAAGTGTCTCACTAACCAATGAGTGGGCTTCACTCAATCTTAAAGAT 3227
 DB 3075 GGGCTGAAGTGAAGTGTCTCACTAACCAATGAGTGGGCTTCACTCAATCTTAAAGAT 3134
 QY 3228 AAAAGTACTATTAATTAACCTTATTAACCAATGAGTGGGCTTCACTCAATCTTAAAGAT 3287
 DB 3135 AAAAGTACTATTAATTAACCTTATTAACCAATGAGTGGGCTTCACTCAATCTTAAAGAT 3194
 QY 3288 TTTGGCAGGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3347
 DB 3195 TTTGGCAGGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3254
 QY 3348 AAAAACAATTTTAAAGATCTTTTAAACAAAAGGCAATGATGATGATGATGATGATGATGATG 3407
 DB 3255 AAAAACAATTTTAAAGATCTTTTAAACAAAAGGCAATGATGATGATGATGATGATGATGATG 3314
 QY 3408 CATATTTATTTTGTGTTGGAATCTGGGACATGATCTAATTTGTATTAATAATTAATGAT 3467

Db 3315 TACTATTATTTCGTTGGACCTGGACATGATCTATTGTTATATAAATAAATGAT 3374

QY 3468 GT 3469
||
Db 3375 GT 3376

RESULT 3
AF387507 3421 bp mRNA linear PRI 01-JUL-2001
LOCUS Homo sapiens PIWI protein mRNA, complete cds.
DEFINITION AF387507
VERSION AF387507.1 GI:14579644
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3421)
AUTHORS Sha, J.H.
TITLE Cloning and identification of human piwi protein related to testis development
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3421)
AUTHORS Sha, J.H., Li, J.M. and Zhou, Z.M.
TITLE Direct Submission
JOURNAL Submitted (30-MAY-2001) Key Lab of Reproductive Medicine, Nanjing Medical University, 140 Han Zhong Road, Nanjing Medical University, Jiangsu 210029, China
FEATURES
Location/Qualifiers
1..3421
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue type="testis"
/dev stage="adult"
64..2649
/note="higher expression in adult testis than 6 month old embryo testis"
/codon start=1
/product="PIWI protein"
/protein_id="AAK69348.1"
/db_xref="GI:14579645"
/translation="MTGRARARARGARQETAQLVGSTASQOPGYIOPRPPPAEG ELPRGRQRTAGTAKSGLQISAFQBSLAERGRDRFDHLDGNTRQNLQHVKE SHFGSGLIVRLSTNFRILTSFQWALYQYHIDYNPLMEARLSALLFOHEDLIGKC HAPDGTILFLPKLOQKTEVFSTRNGEDVRITITLNLPTTSPCLQFVNIIFRR LKIMLQQTGRNYPNDPIDIPSHRLVWPGFTTSILOVENSIMLCTDVSXKLSR ETVLDFMFPYHQTEBHKFOQVSKELIGLVLT.KYNNKTYRVDIDWDQPKSTFKK ADGSEVSFLFYRKQYNOEITDLKQPLVSPQKRRRGGGPGFAMLIPELCYLTGL TDKWRNDFNMKDLAVHRLTPEQREVRLIDYIHKNDNVQRELDRDWGLSPDSNLL SFGSRLIQTEKHOGKTEDPYQFADMSKETRGAPLISVKPLDNWLLIYLRNRYEA NSLIQNLKVTTPANGMKRAIMIEVDDRTEAYLRVLOQKVADTQIVVCLLSNRKD KYDAIKKYLCTDPTPSQCVARTLGKQTVMAITKIALQWCKMGLKVRVDIPLK LVMI VGIDCHMDMTAGRSIAGFVAS INEGTRWFSRCIFODRGOELVDGLKVLQNA LRAMNSCYMPSRIIVYRDGVDGQKATLVNVEPDLCKLSIGRGYAPRLFIYV KKRVTTRFPAQSGSRQLNPGLPTVIDVETRPBWDFIYVQAVRSVSPVTHNVIV DNSGLKPDHILQRLTYLKLCHYYNWEVIRVPAPQYAHKUAFLVQSIHREPNSLSN RLYYL"

polyA_signal 3354..3359
BASE COUNT 1071 a 661 c 767 g 922 t
ORIGIN

Query Match 96.1%; Score 3337.6; DB 9; Length 3421;
Query Local Similarity 99.6%; Pred. No. 0;
Matches 3354; Conservative 0; Mismatches 12; Indels 1; Gaps 1;
QY 104 GCAGAGTCTGAGTGTCAAGGACAGGACTAGGGCGAGGGCGAGCGGTCCAGAAATAGAAA 163
Db 1 GGGGGGGCTGAGGTGCAAGGACAGGACTAGGGCGAGGGCGAGCGGTCCAGAAATAGAAA 60

QY 164 ACAATGACTGGGAGAGCCCGAGCCAGAGCCAGAGAGGGCCCGGGTTCAGGAGACAGCG 223
Db 61 ACAATGACTGGGAGAGCCCGAGCCAGAGCCAGAGAGGGCCCGGGTTCAGGAGACAGCG 120

QY 224 CAGCTGGTGGGCTCCACTGCCAGTGCAGCAACCTGTTTATATTTCAGCCTTAGCCTCAGCCG 283
Db 121 CAGCTGGTGGGCTCCACTGCCAGTGCAGCAACCTGTTTATATTTCAGCCTTAGCCTCAGCCG 180

QY 284 CCACACAGCAGAGGGGGAATTATTTCGGCGTGGACGGCAGAGAGGAAACAGCAGGAGGAACA 343
Db 181 CCACACAGCAGAGGGGGAATTATTTCGGCGTGGACGGCAGAGAGGAAACAGCAGGAGGAACA 240

QY 344 GCCAGTGTCAAGAGCTCCAGATATCTCTGGATTTCAGGAGTTTATCGTAGCAGAGAGA 403
Db 241 GCCAAGTGTCAAGAGCTCCAGATATCTCTGGATTTCAGGAGTTTATCGTAGCAGAGAGA 300

QY 404 GGAGGTCTGCTGAGAGATTTTCATGATCTTGGTGTGAATCAAGCGAGAACCTTAGACCAT 463
Db 301 GGAGGTCTGCTGAGAGATTTTCATGATCTTGGTGTGAATCAAGCGAGAACCTTAGACCAT 360

QY 464 GTTAAAGAAATCAAAAACAGGTTCTTCAGGCAATATAGTAAGGTTTAAGCACTAACCATTTT 523
Db 361 GTTAAAGAAATCAAAAACAGGTTCTTCAGGCAATATAGTAAGGTTTAAGCACTAACCATTTT 420

QY 524 CGGCTGCATCCCGTCCCGAGTGGCCCTTATATCATGATCACATTCAGTATACCACTG 583
Db 421 CGGCTGCATCCCGTCCCGAGTGGCCCTTATATCATGATCACATTCAGTATACCACTG 480

QY 584 ATGGAAGCCAGAGACTCCGTTCACTCTTCTTTTCAACACCAAGATCTAATTTGAAAG 643
Db 481 ATGGAAGCCAGAGACTCCGTTCACTCTTCTTTTCAACACCAAGATCTAATTTGAAAG 540

QY 644 TGTCACTCTTGTGAGAACGATATTTTACCTTAAAGACTACAGCAAAAGGTTACT 703
Db 541 TGCCATCTCTTGTGAGAACGATATTTTACCTTAAAGACTACAGCAAAAGGTTACT 600

QY 704 GAAGTCTTTAGTAAGACCCGGAATGAGAGAGTGTGAGGATAACGATCACTTTTAACAAT 763
Db 601 GAAGTCTTTAGTAAGACCCGGAATGAGAGAGTGTGAGGATAACGATCACTTTTAACAAT 660

QY 764 GAATCTCCACTCATCAACCACTCTGTTTGCAGTTCTATAATTTTTCAGGAGGCTT 823
Db 661 GAATCTCCACTCATCAACCACTCTGTTTGCAGTTCTATAATTTTTCAGGAGGCTT 720

QY 824 TTGAAATCATGAATTTGCAACAAATTTGACGAAATATTATAACCCAAATGACCCAAAT 883
Db 721 TTGAAATCATGAATTTGCAACAAATTTGACGAAATATTATAACCCAAATGACCCAAAT 780

QY 884 GATATTTCAAGTCAAGGTTGAGTTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 943
Db 781 GATATTTCAAGTCAAGGTTGAGTTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 840

QY 944 GAAAAACAGCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 1003
Db 841 GAAAAACAGCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 900

QY 1004 TTGGATTTTCACTGTTCAACTTTTATCATCAGACAGAGAACATAAATTTTCAAGAACAAAGTT 1063
Db 901 TTGGATTTTCACTGTTCAACTTTTATCATCAGACAGAGAACATAAATTTTCAAGAACAAAGTT 960

QY 1064 TCCAAAGAAATTAAGTTAGTTTGTCTTACCAAGTATTAACATTAAGACATACAGAGTG 1123
Db 961 TCCAAAGAAATTAAGTTAGTTTGTCTTACCAAGTATTAACATTAAGACATACAGAGTG 1020

QY 1124 GATGATATTGACTGGGACCCAGAAATCCCAAGAGCACCCTTAAGAGCCGCGGCTCTGGG 1183
Db 1021 GATGATATTGACTGGGACCCAGAAATCCCAAGAGCACCCTTAAGAGCCGCGGCTCTGAA 1080

QY 1184 GTTCAGTCTTTAGAAATCTACAGGAAGCAATATCAACCAAGAGATCACCGACTTTGAAGCAG 1243
Db 1081 GTTCAGTCTTTAGAAATCTACAGGAAGCAATATCAACCAAGAGATCACCGACTTTGAAGCAG 1140

QY 1244 CCTGTCTTGGTTCAGCAGCCCAAGAGAGGGGGCCCTGGGGGACACTGCCAGGCGCT 1303
Db 1141 CCTGTCTTGGTTCAGCAGCCCAAGAGAGGGGGCCCTGGGGGACACTGCCAGGCGCT 1200

Qy	1304	GCACATGCTATTCGAGAGCTCGCTATCTTTCACGCTCTAACCTGAATAATTCGGTATGAT	1363
Db	1201	GCCATGCTCATATTCGAGAGCTCGCTATCTTTCACGCTCTAACCTGAATAATTCGGTATGAT	1266
Qy	1364	TTTAAAGTATATAAAGACTTACGCGCTCATATACAAAGCTAACCTGCAGAGAAAGGACGGT	1423
Db	1261	TTTAAAGTATATAAAGACTTACGCGCTCATATACAAAGCTAACCTGCAGAGAAAGGACGGT	1320
Qy	1424	GAAGTGGAGCATCTATTGATTACATTCATTAACATATATGTTCAAGAGAGCTTCGA	1483
Db	1321	GAAGTGGAGCATCTATTGATTACATTCATTAACATATATGTTCAAGAGAGCTTCGA	1380
Qy	1484	GACTGGGGTTTGAAGTTTGATTCCACATCTACTGCTTCTCAGAGAAATTTTGCAACA	1543
Db	1381	GACTGGGGTTTGAAGTTTGATTCCACATCTACTGCTTCTCAGAGAAATTTTGCAACA	1440
Qy	1544	GAAGAAGTTCACCAAGGTGAAAAAACAATTGATTACATCCCAATTTGCAGATGGTCC	1603
Db	1441	GAAGAAGTTCACCAAGGTGAAAAAACAATTGATTACATCCCAATTTGCAGATGGTCC	1500
Qy	1604	AAAAGAAACAGAGGTGCACATTAATTAAGTGTTAAGCCACTAGATACTGGCTGTGATC	1663
Db	1501	AAAAGAAACAGAGGTGCACATTAATTAAGTGTTAAGCCACTAGATACTGGCTGTGATC	1560
Qy	1664	TATTAAGCGAGAAATTAATTAAGAGCCCAATTCATGATACAAATACTTAATTAAGTTACA	1723
Db	1561	TATTAAGCGAGAAATTAATTAAGAGCCCAATTCATGATACAAATACTTAATTAAGTTACA	1620
Qy	1724	CCAGCCATGCGCATGCAAAATGAGAAAAAGCAATAATGAAGTGGATGACAGAACGAA	1783
Db	1621	CCAGCCATGCGCATGCAAAATGAGAAAAAGCAATAATTAATGAAGTGGATGACAGAACGAA	1680
Qy	1784	GCTTACTTAAGAGTCTTACAGCAAAAGGTCAAGCAGACACCCAGATAGTGTCTGTG	1843
Db	1681	GCTTACTTAAGAGTCTTACAGCAAAAGGTCAAGCAGACACCCAGATAGTGTCTGTG	1740
Qy	1844	TTTGTAAGTATTCGGAAGGACAAATACGATGCTTTAAAAAATACCTGTGTACAGATGTC	1903
Db	1741	TTTGTAAGTATTCGGAAGGACAAATACGATGCTTTAAAAAATACCTGTGTACAGATGTC	1800
Qy	1904	CCATACCCCAAGTACAGTGTGTGTGGCCCGACCTTAAGGCAACAGCAACCTGTACATGGCC	1963
Db	1801	CCATACCCCAAGTACAGTGTGTGTGGCCCGACCTTAAGGCAACAGCAACCTGTACATGGCC	1860
Qy	1964	ATTGCTACAAAGATTGCTCCCTACAGATGAACCTGCAAGATGGAGAGAGCTCTGGAGGGTG	2023
Db	1861	ATTGCTACAAAGATTGCTCCCTACAGATGAACCTGCAAGATGGAGAGAGCTCTGGAGGGTG	1920
Qy	2024	GACATCCCCCTGAAGCTCGTAGATTCGTTGGCATGCTGATTTGTAACATGACATGACAGCT	2083
Db	1921	GACATCCCCCTGAAGCTCGTAGATTCGTTGGCATGCTGATTTGTAACATGACATGACAGCT	1980
Qy	2084	GGGGGAGAGTCAATCGCAGGATTTGTTGCGACATCAATGAAGGATGACCCCGTGGTTC	2143
Db	1981	GGGGGAGAGTCAATCGCAGGATTTGTTGCGACATCAATGAAGGATGACCCCGTGGTTC	2040
Qy	2144	TCACGCTGCATATTTACAGATACAGACAGAGCTGTAGATGGGCTCAAAAGCTGCGCTG	2203
Db	2041	TCACGCTGCATATTTACAGATACAGACAGAGCTGTAGATGGGCTCAAAAGCTGCGCTG	2100
Qy	2204	CAACGGGCTCTGAGGGCTTGGAATAGCTGCAATGATACATGCCAGCCGATCATCGTG	2263
Db	2101	CAACGGGCTCTGAGGGCTTGGAATAGCTGCAATGATACATGCCAGCCGATCATCGTG	2160
Qy	2264	TACCGCATGCGCTAGAGACGGCCAGCTGGAABAAACATGCGTGAACCTACGAAGTCCACAG	2323
Db	2161	TACCGCATGCGCTAGAGACGGCCAGCTGGAABAAACATGCGTGAACCTACGAAGTCCACAG	2220
Qy	2324	TTTTTGGATTGTCTAAATATCAATTGTGTAAGTTACAACCTAGANTAAACGTAAATTTGG	2383
Db	2221	TTTTTGGATTGTCTAAATATCAATTGTGTAAGTTACAACCTAGANTAAACGTAAATTTGG	2280
Qy	2384	GTCAGAAAAAGATGAACCCAGATTTTTTGTCTAGTCTGAGAGAGACTTCAGAAATCCA	2443

Db	2281	GTGAAGAAAAGGTGAACCAAGATTTTTGTCTAGTCTGGAGGAACATTCAGAAATCCA	224
Qy	2444	CTTCTCGGAACAGTTATTTGATGTAAAGTTACCAACAGATGTATGACCTTTTATC	250
Db	2341	CTTCTCGGAACAGTTATTTGATGTAAAGTTACCAAGATGTATGACCTTTTATC	240
Qy	2504	GTGAGCCAGGCTGTGAGAAAGTGTGTGTTCCTCCACAATTAACAATGTACATCTAGAC	256
Db	2401	GTGAGCCAGGCTGTGAGAAAGTGTGTGTTCCTCCACAATTAACAATGTACATCTAGAC	246
Qy	2564	AACAGCGGCTTGAAAGCCAGACCAATACAGGCTTTGACTTAACAGCTGTGCCACATCTAT	262
Db	2461	AACAGCGGCTTGAAAGCCAGACCAATACAGGCTTTGACTTAACAGCTGTGCCACATCTAT	252
Qy	2624	TGCACCTGGCCAGGTGATTCGTGTTCCTGCTCTTGGCAATACGCCCAAGCTGGCT	268
Db	2521	TGCACCTGGCCAGGTGATTCGTGTTCCTGCTCTTGGCAATACGCCCAAGCTGGCT	258
Qy	2684	TTTCTGTGTGGCCAGAGTATTCAGAGAGGCAAAATGTCTCACTGTCAAAACCGCCTTAC	274
Db	2581	TTTCTGTGTGGCCAGAGTATTCAGAGAGGCAAAATGTCTCACTGTCAAAACCGCCTTAC	264
Qy	2744	TACCTCTAACCTGCGAAGAACGATGACCGGCTTTCTTTTGAATGACTTTGGGATTT	280
Db	2641	TACCTCTAACCTGCGAAGAACGATGACCGGCTTTCTTTTGAATGACTTTGGGATTT	2700
Qy	2804	TTTTAAGCTTTTATTTACTTTTTTTTAACTGTAATCTTTCTGGAGGAACCTTGGGAAG	286
Db	2701	TTTTAAGCTTTTATTTACTTTTTTTTAACTGTAATCTTTCTGGAGGAACCTTGGGAAG	276
Qy	2864	GGATTAGGAGATCTAGCAATTTTATTTCTACATTTGCTATTCACCGGCTCTTATTTAT	292
Db	2761	GGATTAGGAGATCTAGCAATTTTATTTCTACATTTGCTATTCACCGGCTCTTATTTAT	282
Qy	2924	ATGTAAATTTAAATTTTATTAATTTTATCTTCTGTGTCTCATAGATATTTTGTAGCAT	298
Db	2821	ATGTAAATTTAAATTTTATTAATTTTATCTTCTGTGTCTCATAGATATTTTGTAGCAT	288
Qy	2984	TTTTTGTGTTATTTTGAAGAAATGTGAATGAATCTGTGTGTATTAACAGACTCTCT	304
Db	2881	TTTTTGTGTTATTTTGAAGAAATGTGAATGAATCTGTGTGTATTAACAGACTCTCT	294
Qy	3044	GAGAGTATTTGAAAATGTGTGTGAGATTTTACTTAAACGTACTTTGAGAGTGACAAATC	310
Db	2941	GAGAGTATTTGAAAATGTGTGTGAGATTTTACTTAAACGTACTTTGAGAGTGACAAATC	300
Qy	3104	CTACTTAATTAACCAATATTTAACTTTATTTTGTGAGATCTGCTGTTTGAATTTAAGAGAT	316
Db	3001	CTACTTAATTAACCAATATTTAACTTTATTTTGTGAGATCTGCTGTTTGAATTTAAGAGAT	306
Qy	3164	AAGAGCGTAAAGTAGAGTCTCACTACACACATAGTGTGGGTTTCAAGTCAATCTTA	322
Db	3061	AAGAGCGTAAAGTAGAGTCTCACTACACACATAGTGTGGGTTTCAAGTCAATCTTA	312
Qy	3224	AGATTAAGGTATCTATTAATATTAACCTTAACAAGATACAGAGAAAATATGCTTGATTT	328
Db	3121	AGATTAAGGTATCTATTAATATTAACCTTAACAAGATACAGAGAAAATATGCTTGATTT	318
Qy	3284	TTATTTGGCAGGGGGCTAGGTGTATGGAGGT-AAAAAAAACATGGAATTTTATTAAT	334
Db	3181	TTATTTGGCAGGGGGCTAGGTGTATGGAGGTAAAAAAAACATGGAATTTTATTAAT	324
Qy	3343	TGTCCAAAGAACATTTTAAAGCTCTTTAACAAAAAAGCGCATGAGTAATCTCTATATT	340
Db	3341	TGTCCAAAGAACATTTTAAAGCTCTTTAACAAAAAAGCGCATGAGTAATCTCTATATT	330
Qy	3403	AACATCACTATTAATTTTGTGTGAACTGGGACATGATCTATTTGTATTAATTAATAA	346
Db	3301	AACATTACTATTAATTTTGTGTGAACTGGGACATGATCTATTTGTATTAATTAATAA	336
Qy	3463	TTGATGT 3469	

RESULT 6	AB032604	3910 bp, mRNA	linear	ROD 04-OCT-2001
LOCUS	AB032604			
DEFINITION	Mus musculus mRNA for MIM1, complete cds.			
ACCESSION	AB032604			
VERSION	AB032604.1	GI:7416110		
KEYWORDS	MIM1 (pim1).			
SOURCE	Mus musculus	tissue_lib:testis cDNA to mRNA.		
ORGANISM	Mus musculus			

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL

1 Kuramochi-Miyagawa, S., Kimura, T., Yomogida, K., Kuroiwa, A., Tadokoro, Y., Fujita, Y., Sato, M., Matsuda, Y. and Nakano, T. Two mouse p1w1-related genes: miw and mli Mech. Dev. 108 (1-2), 121-133 (2001)
21463379
2 (bases 1 to 3910)
Miyagawa, S.K. and Nakano, T.
Direct Submission
Submitted (21-SEP-1999) Satomi K Miyagawa, Research Institute for Microbial Diseases, Department of Molecular Cell Biology, Yamadoko 3-1, Suita, Osaka 565-0871, Japan
(E-mail: smiyagawa@iken.osaka-u.ac.jp, Tel.:81-6-6879-8362, Fax:81-6-6879-8362)

```

FEATURES
source
Location/Qualifiers
1..3910
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="5F distal-G2 proximal"
/tissue_1ib="testis"
1..3910
/gene="miwi"
156..2744
/gene="miwi"
/note="piwi family"
/codon_start=1
/product="MIMI"
/protein_id="BAA93705.1"
/db_xref="GI:7416111"
/translacion="MTGARARARARGARGAQETVQHVGAANAASQPGCITPPRPQOSPTET
DIVAGRGQGVAVGATSKSQEOLQISAGFQELSLARSGGRBPHLGVNTRQNLPHVIA
EKTQSGGIIYKLSHTKHPFLTRSPQWALVYOHIDYNNPLMEARLSALFQHEDLIG
CHAPGDTLFLPKRLQHKVTEVSOTRNGEHRITTLTNELPTPSCTLOPYNIIFFI
RLKTKMLNLOQIGRANTYNPSDPIIDIPNHRIVYMPGFTTSLQYENNMIMCTVSHYVL
STVDLVEFNLIVQQTBEHKFPQEQVSLGLGLVLTKYNKTKTRVDDIMDDQPKSTVF
KADGESEVFLLEYVRQYQOEITDLQKPYLVSQPKRRRGCGTLPGPAMIIPBLCVLYT
LIDKRNDENWAKDOLAVHTRLPEQGRREVGLIDYIHDDVQRELRBMGSLPDSNIN
LSFSRRIIOSEKIHOGKTFEDYNPOPADMSKEFRGAPLISVPLDNMLLIYTRRYVE
ANSILONLFEKYPANGIOKMKALIMEVDRTEAYALALOOQXVTSPOQIVYVCLLSNR
DIAMVIAKTLCTDCPTSPQCVAVRTGKOQYMAIKTALQNCNGGELVRVDMP
KLYMTVDICHDYTTAGRSIAGFVASINEGMTWRSRCVFPDRQGEVLGKVCLOQ

```

QY	164	ACAATGACTGGAGAGCCCGAGCCAGAGCCAGAGGAAGGGCCCGGCTCAGAGAGCAGCG	223
Db	153	AAAATGACTGGCGCGAGCCCGAGCTCGGGCCCGCGCAGAGGACAGAGGCTCAGAGACGGTG	212
QY	224	CAGCTGGTGGGCTCCACTGCGCAGTCAGGAACTCTGGTTATTTACGCTTAGCGCTCAGCGC	283
Db	213	CAGCAATTTGGGGCTGTCTGCGAGCCAGCACTTGGGTACATCCCACTCGAGACTTAAACAG	272
QY	284	CCACGACGAGAGGGGGAAATTATTTGGCCGTGACGGCAGAGAGGAA---CAGCAGAGAGA	340
Db	273	TCCCCACAGAGGGGGACTTGGTTGGCCGAGGACAGACAGAGGGAGTGTAGTCCGAGCC	332
QY	341	ACAGCCAATGTCACAGAGACTCAGATATCTGTGTGAATTCACGAGTTATTCGTAGCAGAG	400
Db	333	ACATCCAGATCAAGAAACTGACAGATCTGAGTGGGTTTCAGAGACTGTCACTGGCAGAG	392
QY	401	AGAGAGGTCGTCTGTAAGATTTTTCATGATCTTGGTGTGAATTAAGGACAGAACTTAGC	460
Db	393	AGAGAGAGGGCGTCCGCGAGACTTCATGACCTTGGTGTGAACACGACAGAACTTAGC	452
QY	461	CATGTTAAAGATCAAAAACAGGTTCTTCAGGCATTTATATGTAAGTTAAGCACTAACAT	520
Db	453	CATGTCAAAGATCAAAAGCAGGCTCTCTGCAATCATTTGTAACCTGAGCAACAAACAC	512
QY	521	TTCCGGCTGACATCCCGTCCCGAGTGGGCTTATATCAGTATCACTTGACTATTAACCA	580
Db	513	TTCCGGCTGACCTCGCGCCCAAGTGGGCGCTGTATCAGTACCAATGACCTAACATCC	572
QY	581	CTGATGGAAGCCAGAAAGACTCCGTTACGCTCTCTTTTCAACGAGAGATCTAATTGGA	640
Db	573	CTGATGGAAGCCAGAAAGGCTTCGCTCCGCACTGCTCTTCACACAGAAAGCTCAATTGGA	632
QY	641	AGTGTCAATGCTTTGATGAGAAAGATATTTATACCTAAGAAAGCTAAGCAAAAGTT	700
Db	633	AGTGTCAATGCTTTGATGAGAAAGATATTTATTTTACCTAAGAAAGCTAAGCAAAAGTT	692
QY	701	ACTGAAGTTTATGTAAGACCCGGAATGAGAGATGTGAGATTAAGATCACTTTTACA	760
Db	693	ACAGAAAGTATCAGTCAGACTCGGAATGGGGAACGTAAGATCACTATCACTCCGAC	752
QY	761	AATGAACCTTCCACCTACATCAACCACTTGTGTGCACTTATATATATTTTTCAGAGG	820
Db	753	AACGAGCTCCGCGCACCTCGCCACCTGCGCAGCTTATTAACATCACTTTCAGAGG	812
QY	821	CTTTGAAATCATGAATTTGCAACAATTTGACAGAAATTTATTAACCAATGAACCA	880
Db	813	CTCTTGAATCATGAATTTGCAACAATTTGACAGAAATTTATTAACATCAAGTGAACCG	872
QY	881	ATTGATATTTCAAGTCACAGGTTGTGATTTGGCTGTCACTACCTTCATCTTTGAG	940
Db	873	ATTGATATTTCAAAACACAGGTTGTGATCTGGCCGCGCTTCAACCACTTCATCTTTGAG	932
QY	941	TATGAAAGACGATCAATGCTGCACTGACGTGACGTTAAGCATTAAGTCTTGAAGTGAAGCT	1000
Db	933	TATGAAAGACGATCAATGCTGCACTGACGTCGTCAGCCCAAGTGTCCGCAACGAGACT	992
QY	1001	GTTTGGATTTCAATGTTCAACTTTATATCATCAGACAGAAACATTAATTTCAAGAACAA	1060
Db	993	GTCCTGATCTTCAATGTTCAATCATATCCAGACAGAGAGACACAAAGTTCAGAGGACAA	1052
QY	1061	GTTTCAAGAAATTAATAGTTTGTAGTTGTTCTTACCAAGTATTAACCTTAAGCATACAG	1120

Db 1053 GTGTGAAGGAGCTCATAGGCTTCATCGTTCTCACCAAGTACAAATACAAAGACCTACCGG 1112
 Qy 1121 GTGATGATATTTGACTGGGACGAGATCCAGAGGACCTTTTAAAGAACGCCGAGCGCTCT 1180
 Db 1113 GTGATGACATTTGACTGGGACGAGATCCAAAGAGCACCTTTCAAGAGGCGGATGGCTCG 1172
 Qy 1181 GGGGTACAGCTTCTTGAATATCTACAGGAAGCAATACAAACCAAGAGATCACCGACTTTGAAG 1240
 Db 1173 GAGGTACAGTCTCTGGAGTACTACAGGAAGCAATACAAACCAAGAGATCACCGACTTTGAAG 1232
 Qy 1241 CAGCTCTCTTTGGTTCAGCCAGCCCAAGAGAGGCGGCGCTTGGGGGACACTGCCAGGG 1300
 Db 1233 CAGCGGTGCTGTGTAGGACCAACCAAGCGGAGAGAGGCCCCGCGGCGACCTTGCTGGC 1292
 Qy 1301 CTGCGATGCTCATCTCTGAGCTCTGCTATCTTACAGGTCTAACTGATAAATATGCGTAAAT 1360
 Db 1293 CCAGCTATGCTCATCTCCCTGAACCTGCTATCTCACAGGCTGACTGATAAATATGCGCAAT 1352
 Qy 1361 GATTTTAACTGTATGAAGACTTAGCCGTTTCATACAGACTACTCCAGAGCAAGGCGAG 1420
 Db 1353 GATTTTCAATGTATGAAGGACCTGGCAGTGCACACGCGGCTGACCCCTGAGCGGCGAG 1412
 Qy 1421 CGTGAAGTGGGAGCACTCATTTGATTAATTCATTAATAAAACGATTAATGTTCAAAGGAGCTT 1480
 Db 1413 CGGAGGTGGCGCGCTCATCGACTACATCCACAAGGATGACAAATGTGCAGAGAGACTT 1472
 Qy 1481 CGAGACTGGGTTTGAAGCTTTGATTTCAACTTACTGCTCTTCTCAGAGAGAAATTTGAAA 1540
 Db 1473 CGAGACTGGGCGCTGAGCTTCGACTCAAACTTGTCTCTCTCTGGAAGAAATCTTACAA 1532
 Qy 1541 ACAGAAAGATTCACCAAGGTGGAAACAAATTTGATTACAAATCCACAATTTGCAGATCG 1600
 Db 1533 TCTGAGAAGATTCACCAAGGCGGAAAGAGCTTTGATTACAAACCAATTTGCAGACTGG 1592
 Qy 1601 TCCAAAGAAAACAAGAGGTGCACCAATTAATTTAGTGTAAAGCACTAGATAACTGGCTGTG 1660
 Db 1593 TCCAAAGAAAACAAGAGCGCGCTGATCAGCGTGAAGCCATTGATTAACCTGGCTGTG 1652
 Qy 1661 ATCTATACGGAAGAAATATTAAGCAGAGCAATTCATTGATACAAATCTATTTAAAGTT 1720
 Db 1653 ATCTATACCGGAGGAATATTAAGCAGAGCAATCTCACTGATACAGAACTGTTCAAAGTG 1712
 Qy 1721 ACACGAGCACTGGCATGCAAAATGAGAAAGCAATTAATGATTGAAGTGATGACAGAACT 1780
 Db 1713 ACTCCAGCACTGGCATGCAAGTGAAGAGCAATCATGATCAGGTGATGACAGAACT 1772
 Qy 1781 GAAGCCTACTTAAGTCTTACAGCAAAAGGTTCACAGCAGACACCCAGATAGTTGTCTGT 1840
 Db 1773 GAAGCTTATCTGAGAGCTTTCAGCAGAGAGGTGACGCTCAGACACTCAGATAGTTGTCTGT 1832
 Qy 1841 CTGTTGTCAAGTAATCGGAAGGCAAAATAGGATGCTATTAAAAATACCTGTGTACAGAT 1900
 Db 1833 CTCTTGTCAAGTAATCGGAAGGCAAAATATGATGCCATCAAGAAGTACTTGTGTACAGAC 1892
 Qy 1901 TGCCTTACCCCAAGTCAAGTGTGTGTGGCGGCAACCTTAGGCAAAACAGCAAACTGTCTATG 1960
 Db 1893 TGCCCAACCCCAAGTCAAGTGTGTGTGGCGGCAACCTTAGGCAAAACAGCAAACTGTCTATG 1952
 Qy 1961 GCAATTTGCTACAAAGATTCCTTACAGATGAACCTGCAAGATGGGAGGAGCTCTGGAGG 2020
 Db 1953 GCAATTTGCCAACCAAGATTCCTTACAGATGAACCTGCAAGATGGGAGGAGCTCTGGCGG 2012
 Qy 2021 GTGGACATCCCTTGAAGTCTGTGATGATCGTTGGCATGATGTTTACCATCAGATGACA 2080
 Db 2013 GTGGACATCCCTTGAAGTCTGTGATGATCGTTGGCATGATGTTTACCATCAGATGACA 2072
 Qy 2081 GTGGCGGAGGTCAATTCGAGGATTTGTTGCCAGCATCAATGAAGGGATGACCCGCTGG 2140
 Db 2073 GCTGGCGGAGGTCCATCGAGGATTCGTGCCAGCATCAATGAAGGGATGACCCGCTGG 2132
 Qy 2141 TTCTACGCTGCATATTTTCCAGATAGAGACAGGAGCTGGTATAGTGGGCTCAAAATCTGCG 2200
 Db 2133 TTCTCCGCTGCGTCTTTTCAGGACCGCGGCGAGGAGCTGGTGGATGGTCTCAAGGTGTGC 2192

Qy 2201 CTGCAAGCGGCTCTGAGGGCTTGGAAATAGCTGCAATGAGTACATGCCCAGCGGATCATC 2260
 Db 2193 TTGCAAGCTGCTCTGAGGGCTTGGAGTGGCTGCAATGAATACATGCCCAGCGTGTCTATC 2252
 Qy 2261 GTGTACCCGCGATGCGTAGGAGACGGCCAGCTGAAACACATGCTGGAACCTACGAAGTGCCA 2320
 Db 2253 GTGTACCCGAGACGGTGTGGGGACGGGCGAGCTGAAGACCCCTGGTCAATATGAGGTCCCA 2312
 Qy 2321 CAGTTTTTGGATTTGCTAAAAATCCATTTGTTAGAGTTTACAAACCTAGANTAAACGGTAATT 2380
 Db 2313 CAGTTCTTAGATTCCTCAAGTCACTCGGAGAGGTTTACAAACCAAGACTGACTGTATATC 2372
 Qy 2381 GTGTGAAGAAAAGAGTGAACACACAGATTTTTTGTCTCAGTCTGGAGGAAGACTTTCAGAAAT 2440
 Db 2373 GTGTGAAGAAAAGCGTGTCAATGCCAGGTTTTTGTCTCAGTCTGGGGGAAGACTTTCAGAAAC 2432
 Qy 2441 CCACCTCTCTGAAACAGTTTATGATGTAGAGTTTACAGAGCCAGCAATGTTATGACTTTTTT 2500
 Db 2433 CCTCTTCCAGGGACAGTCAATCGATGTGGAAGTCCACAGACAGAGTGTATGACTTTTTT 2492
 Qy 2501 ATCTGTAGCCAGGCTGTGAGAAAGTGTGTGTTTCTCCCAACATTAACAATGTCTATCTAT 2560
 Db 2493 ATCTGTAGCCAGGCAAGTGTGAGAGCGGAGTGTGTCCCAACACACTACATATGTCTAT 2552
 Qy 2561 GACAAAGCGGCTTGAAGCCAGACCAATACAGGCTTGTGACCTTCAAGCTGTGCCAATC 2620
 Db 2553 GACAGCAGTGGCTTGAAGCCGCGACCAATCCAGCGGTGACATACAAAGCTCTGCCAGGTG 2612
 Qy 2621 TATTACAACTGGCCAGGCTGTCATTCGTGTTCTCTCTGTCAGTACGCCAGTACGCCCAAGCTG 2680
 Db 2613 TACTATTAATTTGGCTTGGAGTCACTCCGAGTCCCTGACCTTGGCAGTATGCACAAAGCTG 2672
 Qy 2681 GCTTTTCTTGTGGCCAGAGTATTTCAGAGAGCGCAAAATCTGTCACTGTCAAAACCGGCTT 2740
 Db 2673 GCCTTCTCTGTGGGCCAGAGCAATCCACAGAGCGCAAACTCTCCCTGTCCAACCGCTC 2732
 Qy 2741 TACTACCTCTAACTGCGAG 2759
 Db 2733 TACTACCTCTAAACAGCAG 2751

RESULT 7
 AF438405 Mus musculus MIWI mRNA, complete cds. linear ROD 13-NOV-2001
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Mus musculus.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 4064)
 Deng, W. and Lin, H.
 Miwi, a murine homolog of piwi, encodes a cytoplasmic protein
 essential for spermatogenesis
 Unpublished
 2 (bases 1 to 4064)
 Deng, W. and Lin, H.
 Direct Submission
 Submitted (22-OCT-2001) Department of Cell Biology, Duke
 University, Room 412, Nannaline Duke Building, Durham, NC 27710, USA
 Location/Qualifiers
 1..4064
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="5"
 /tissue_type="testis"
 /note="piwi-like protein"
 /codon_start=1
 /product="MIWI"

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

FEATURES
 source
 1..4064
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="5"
 /tissue_type="testis"
 /note="piwi-like protein"
 /codon_start=1
 /product="MIWI"

CDS

```

/protein_id="AAJ31014.1"
/db_xref="GI:16905061"
/translation="MTGRARARARGARAGCTVQVHGAASQCPGTYIPRPOQSPTEG
DLVGRGRQWVATGSKSLOISAGFQESLBERGGRBDPDLVNRONLDHYK
ESKTSGGIIVKLSMNRRLTSRPMWLXOYHIDNPLMRRLRSLLIHOHEILIR
CHAPGTLIPKRLQHKVTEPSPQTRNGEVRITITLTNETPTSTCLOFNITIR
RLKLTMLNQQIGRNTINPSDPIIDPNHLLVWPGPTSLIYENNIMLCTDVSHKVLK
SETVIDEFNLIQOETENKFOVSKELGLVLTKNKKTVRVDIDIMONPKSTPK
KADGSEVFLERYKQYNIETDLKQPVLPYRPRRPGGTLPAMLIPELCYLG
LTKMRNDPFWKDLAVHRLRLEPORQVRLIDYHKNDRVORELRDGLSPDSNL
LSFSGRILOSEKHOGKTPDYNPOFADMSKEBGAPLISYKLDNMLLIYTRNVA
ANSLIOLFKYTPMGIOMKRAMIENDRREARLALQOQVTSPTDITVCLISNRK
DKYDIKRYLCTDCTPSQCYVARTLQOVTMALIKTALQNNCKNGBELMRVDEL
KLAMIVGIDCYHDITPARRSIAGFVASINEBETWFSRQVODRGEIYGLKVCLOA
ALRAMSGNEVMPSRVIVYRDGVDGQKTLVNEVQFDELCKSVGRVPRLTIV
YKRVNARPFPAQSGRLONPLPGYIVDIEVTRPMPYEPFISYASVRSQSVPTNVY
YDSGLKPEDHIOBLRYKLCHYVNMWPGVIRPAPCOYAHKLAFLVQSHIREPLSLIS
NRLYTL"
BASE COUNT      1115 a      982 c      1077 g      890 t
ORIGIN

```

```

Query Match      53.0%; Score 1839.2; DB 10; Length 4064;
Best Local Similarity 82.0%; Pred. No. 0;
Matches 2130; Conservative 0; Mismatches 466; Indels 3; Gaps 1;

```

```

QY 164 ACAATGACTGGAGAGCCCGGACCAAGCCAGAGGAAGGCGCGGTCAAGAGACAGCG 223
DB 188 AAAATGACTGCGGAGCCGAGCTCGGCGCGGAGGAGGACGAGTCAAGAGACGCGTG 247
QY 224 CAGCTGTGGGCTCCAGCTCCAGCTCAGCACTGTGTATTTACGCTTACGCTTACGCG 283
DB 248 CAGCATGTGGGCTGCTCGGAGCAGCACTGTGTATTCCTCCAGCACTTCAGAG 307
QY 284 CCACGAGCAGAGGGGGAATTTATTTGGCGGTGAGCGGAGAGAGAA-----CAGAGAGAGA 340
DB 308 TCCCCCAGAGGGGGAATTTGTGTGGCGGTGAGCGGAGAGAGAGAGAGAGAGAGAG 367
QY 341 ACAGCCAGTCAAGAGCTCCAGATCTGTGTATTTACAGATTTACAGATTTACAGAG 400
DB 368 ACATCCAGTCAAGAGCTCCAGATCTGTGTATTTACAGATTTACAGATTTACAGAG 427
QY 401 AAGAGAGCTGCTGAGAGATTTTATGATCTGTGTATTTACAGATTTACAGAGAGAG 460
DB 428 AAGAGAGGCGGTGCGGAGCTTCATGATCTGTGTATTTACAGAGAGAGAGAGAG 487
QY 461 CATGTTAAAGATCAAAAAAGAGTCTTCAGGATTTATAGTAAAGTAAAGAT 520
DB 488 CATGTTAAAGATCAAAAAAGAGTCTTCAGGATTTATAGTAAAGTAAAGAT 547
QY 521 TTCCGGCTGACATCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 580
DB 548 TTCCGGCTGACATCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 607
QY 581 CTGATGAGAGCCAGAGAGCTCGTTCAGCTCTTTTCAACAGAGAGATTTAAGAG 640
DB 608 CTGATGAGAGCCAGAGAGCTCGTTCAGCTCTTTTCAACAGAGAGATTTAAGAG 667
QY 641 AAGTGTCACTTTTGTATGAGAGATTTATTTTACCTAAAGATCAAGAGAGAGTT 700
DB 668 AAGTGTCACTTTTGTATGAGAGATTTATTTTACCTAAAGATCAAGAGAGAGTT 727
QY 701 ACTGAAAGTTTATGAGAGCCGGAATGAGAGATGAGAGATTAAGATCACTTTAACA 760
DB 728 ACAGAAAGTATTCAGTCAAGCTCGGAATGAGAGATGAGAGATCAAGATCAAGCTG 787
QY 761 AATGAACTTCCAGTCAAGATCAAGATCTGTGTGAGTCTTATATATTTATTTTGAAG 820
DB 788 AATGAACTTCCAGTCAAGATCAAGATCTGTGTGAGTCTTATATATTTATTTTGAAG 847
QY 821 CTTTGAAGATCAAGATTTTGAAGAGAGATTTTGAAGAGATTTTGAAGAGAGAGAG 880
DB 848 CTTTGAAGATCAAGATTTTGAAGAGAGATTTTGAAGAGATTTTGAAGAGAGAGAG 907

```

```

QY 881 ATTGATATTCAGATCAAGAGTGTGATTTGCGCTTCACTTCCATCTTCAAG 940
DB 908 ATTGATATTCAGATCAAGAGTGTGATTTGCGCTTCACTTCCATCTTCAAG 967
QY 941 TATGAAAGAGATCAAGATCTTCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1000
DB 968 TATGAAAGAGATCAAGATCTTCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1027
QY 1001 GTTTGATTTCAATGTTCACTTTTATCATCAGAGAGAGATTAATTTCAAGAGAA 1060
DB 1028 GTCTGATCTTCAATGTTCACTTTTATCATCAGAGAGAGATTAATTTCAAGAGAA 1087
QY 1061 GTTTCAAGAGATTAATGAGTGTGTTGTTTCTTCAAGATTAATTAAGATTAAG 1120
DB 1088 GTGTCAGAGAGATTAATGAGTGTGTTGTTTCTTCAAGATTAATTAAGATTAAG 1147
QY 1121 GTGATTAATTTGATGAGAGCCAGAGATTTCCAGAGAGATTTTGAAGAGAGAG 1180
DB 1148 GTGATTAATTTGATGAGAGCCAGAGATTTCCAGAGAGATTTTGAAGAGAGAG 1207
QY 1181 GGGGTCACTTTCTTGAATCTACAGAGAGATTAATTAAGAGATTAATTAAGAG 1240
DB 1208 GAGTCACTTTCTTGAATCTACAGAGAGATTAATTAAGAGATTAATTAAGAG 1267
QY 1241 CAGCTGTCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1300
DB 1268 CAGCTGTCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1327
QY 1301 CTGCTGATCTTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1360
DB 1328 CAGCTGATCTTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1387
QY 1361 GATTTTAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1420
DB 1388 GATTTTAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1447
QY 1421 CGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1480
DB 1448 CGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1507
QY 1481 CGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1540
DB 1508 CGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1567
QY 1541 ACAGAAAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1600
DB 1568 TCTGAAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1627
QY 1601 TCCAAAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1660
DB 1628 TCCAAAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1687
QY 1661 ATCTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1720
DB 1688 ATCTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1747
QY 1721 ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1780
DB 1748 ACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1807
QY 1781 GAGGCTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1840
DB 1808 GAGGCTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1867
QY 1841 CTGTTGTCAAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1900
DB 1868 CTGTTGTCAAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1927
QY 1901 TGCCCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1960
DB 1928 TGCCCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1987
QY 1961 GCCATTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2020

```

JOURNAL
Unpublished
essential for germ-line stem cell self-renewal

QY 693 AAAAGTTAAGTATTTTAAAGACCGGAAATGAGAGATGAGATTAAGATCA 752
 Db 584 ACCCTGAGCAAGTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 643
 QY 753 CTTTAAATGAGATTTTCAAGTATTAAGATTTTGAAGTATTAAGATTTT 812
 Db 644 CACTGACATGAGTGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 703
 QY 813 TCAAGAGCTTTTGAAGATTAAGATTTTGAAGATTTTGAAGATTTTGAAG 763
 Db 704 TCAAGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTT 763
 QY 873 ATGACCCATTTGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 932
 Db 764 ATGACCCATTTGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAG 923
 QY 933 TCTTCAAGTATTAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTT 992
 Db 824 TCTTCAAGTATTAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTT 883
 QY 993 GTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTT 1052
 Db 884 GTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTT 943
 QY 1053 AAGACATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTT 1112
 Db 944 CTGAGATTTTCAAGATTTTCAAGATTTTCAAGATTTTCAAGATTTT 1003
 QY 1113 CATACAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTT 1172
 Db 1004 CATACAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTT 1063
 QY 1173 ACGGCTCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 1232
 Db 1064 A---TACAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 1120
 QY 1233 ACTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTT 1292
 Db 1121 ATGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTT 1180
 QY 1293 TGCAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTT 1352
 Db 1181 CACCGGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTT 1240
 QY 1353 TGCAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTT 1412
 Db 1241 TGCAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTT 1300
 QY 1413 AAGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTT 1472
 Db 1301 AAGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTT 1360
 QY 1473 GGGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTT 1532
 Db 1361 ATGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTT 1420
 QY 1533 TTTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTT 1592
 Db 1421 TTTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTT 1480
 QY 1593 CAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 1652
 Db 1481 CAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 1540
 QY 1653 GGGTGTATTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 1712
 Db 1541 GGGTGTATTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 1600
 QY 1713 TTTAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 1772
 Db 1601 ACAAAGTGTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 1660

QY 1773 AAGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 1832
 Db 1661 AAGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 1720
 QY 1833 TTTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 1892
 Db 1721 TTTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 1780
 QY 1893 GTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 1952
 Db 1781 GTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 1840
 QY 1953 GTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2012
 Db 1841 GTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 1900
 QY 2013 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2072
 Db 1901 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 1960
 QY 2073 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2132
 Db 1961 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2020
 QY 2133 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2192
 Db 2021 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2080
 QY 2193 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2252
 Db 2081 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2140
 QY 2253 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2312
 Db 2141 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2200
 QY 2313 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2372
 Db 2201 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2260
 QY 2373 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2432
 Db 2261 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2320
 QY 2433 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2492
 Db 2321 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2380
 QY 2493 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2552
 Db 2381 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2440
 QY 2553 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2612
 Db 2441 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2500
 QY 2613 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2672
 Db 2501 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2560
 QY 2673 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2732
 Db 2561 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2620
 QY 2733 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2792
 Db 2621 TCTGAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAGATTTTGAAG 2660

RESULT 9
 AC025837/c AC025837 DNA linear HTG 06-MAY-2001
 LOCUS Homo sapiens clone RP11-200K12, WORKING DRAFT SEQUENCE, 29
 DEFINITION

```

unordered pieces.
ACCESSION AC025837
VERSION AC025837.3 GI:8072609
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 193697)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
TITLE Homo sapiens chromosome, clone RP11-200K12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 193697)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhgaiter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galegan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczyk,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessfay,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced gi:7342026.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5437
Center clone name: 200 K.12
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Assembly: Dye-terminator Big Dye; 100% of reads
Consensus program: Phrap; version 0.960731
Consensus quality: 174594 bases at least Q40
Consensus quality: 183716 bases at least Q30
Consensus quality: 187939 bases at least Q20
Insert size: 183000; agarose-fp
Quality coverage: 3.9 in Q20 bases; agarose-fp
Quality coverage: 3.7 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 29 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1628: contig of 1628 bp in length
* 1629 1728: gap of 100 bp
* 1729 2962: contig of 1234 bp in length
* 2963 3062: gap of 100 bp
*

```

```

* 3063 4989: contig of 1927 bp in length
* 4990 5089: gap of 100 bp
* 5090 5666: contig of 1577 bp in length
* 5667 6766: gap of 100 bp
* 6767 8667: contig of 1901 bp in length
* 8668 8767: gap of 100 bp
* 8768 11259: contig of 2492 bp in length
* 11260 11359: gap of 100 bp
* 11360 13557: contig of 2198 bp in length
* 13558 13657: gap of 100 bp
* 13658 17173: contig of 3516 bp in length
* 17174 17273: gap of 100 bp
* 17274 20060: contig of 2787 bp in length
* 20061 20160: gap of 100 bp
* 20161 23843: contig of 3683 bp in length
* 23844 23943: gap of 100 bp
* 23944 27203: contig of 3260 bp in length
* 27204 27303: gap of 100 bp
* 27304 31592: contig of 4289 bp in length
* 31593 31692: gap of 100 bp
* 31693 36740: contig of 5048 bp in length
* 36741 36840: gap of 100 bp
* 36841 42197: contig of 5357 bp in length
* 42198 42977: gap of 100 bp
* 42978 47652: contig of 5355 bp in length
* 47653 47752: gap of 100 bp
* 47753 52307: contig of 4555 bp in length
* 52308 52407: gap of 100 bp
* 52408 58606: contig of 6199 bp in length
* 58607 58706: gap of 100 bp
* 58707 63584: contig of 6878 bp in length
* 63585 65684: gap of 100 bp
* 65685 72492: contig of 6808 bp in length
* 72493 72592: gap of 100 bp
* 72593 78756: contig of 6164 bp in length
* 78757 78856: gap of 100 bp
* 78857 85474: contig of 6618 bp in length
* 85475 85574: gap of 100 bp
* 85575 91951: contig of 6377 bp in length
* 91952 92051: gap of 100 bp
* 92052 99044: contig of 6993 bp in length
* 99045 99144: gap of 100 bp
* 99145 104884: contig of 5740 bp in length
* 104885 104984: gap of 100 bp
* 104985 112130: contig of 7146 bp in length
* 112131 112230: gap of 100 bp
* 112231 118017: contig of 5787 bp in length
* 118018 118117: gap of 100 bp
* 118118 126319: contig of 8202 bp in length
* 126320 126419: gap of 100 bp
* 126420 137651: contig of 11232 bp in length
* 137652 137751: gap of 100 bp
* 137752 193697: contig of 55946 bp in length.
FEATURES
source
Location/Qualifiers
1..193697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-200K12"
/misc_feature
1..1628
/note="assembly_fragment"
misc_feature
1729..2962
/note="assembly_fragment"
misc_feature
3063..4989
/note="assembly_fragment"
misc_feature
5090..5666
/note="assembly_fragment"
misc_feature
5667..8667
/note="assembly_fragment"
misc_feature
8768..11259
/note="assembly_fragment"
misc_feature
11360..13557
/note="assembly_fragment"

```


	Db	1387122	CTAGCAATTTTATTTCCTAGACATTCGTAATTCACCGGGCTCCTTATTTTAAACGAAAAAATTA	138663
	Oy	2936	AGATTATAATTTTAACTCTTGTTTTCTCATAGATAATTTTGTGACATTTTTCCTTAT	2995
	Db	138662	AGATTATAATTTTATTCCTCTTGTTTTCTCATAGATATTTTGTGAGCATTTTTCCTTAT	138603
	Oy	2996	TTTGAAAATAATGTGATAGAATACTCTTGATATTAATAACAGACTCTCTGAGAGTATTTGA	3055
	Db	138602	TTTGAAAATAATGTGATAGAATACTCTTGATATTAATAACAGACTCTCTGAGAGTATTTGA	138543
	Oy	3056	AATGCTTTGGAGATTTACTTAAACGTAATTTTCAGGAGTAGAGCAAGTCTTACTTATTAAC	3115
	Db	138542	AATGCTTTGGAGATTTACTTAAACGTAATTTTCAGGAGTAGAGCAAGTCTTACTTATTAAC	138483
	Oy	3116	CTATTATTAATTTTATTTTGTGAGATACCTGTTTTGAATTTAAAGAGATAGAGCGGTAA	3175
	Db	138482	CTATTATTAATTTTATTTTGTGAGATACCTGTTTTGAATTTAAAGAGATAGAGCGGTAA	138423
	Oy	3176	GTAAGATGCTCACATCAACATAGGTGGGGTTTCAGCTCATATCTTAAAGATTAAGGTA	3235
	Db	138422	GTAAGATGCTCACATCAACATAGGTGGGGTTTCAGCTCATATCTTAAAGATTAAGGTA	138363
	Oy	3236	CTATTATTAATCTTATTAACAAGATTAACAGAGAAATATGCTGATTTTATTTTGCAGG	3295
	Db	138362	CTATTATTAATCTTATTAACAAGATTAACAGAGAAATATGCTGATTTTATTTTGCAGG	138303
	Oy	3296	GGGGCTAGGCTGTATGAGGAGTAAAAAAAAAACATTTGAAATTTTAAATGTCGAAGAAC	3355
	Db	138302	GGGGCTAGGCTGTATGAGGAGTAAAAAAAAAACATTTGAAATTTTAAATGTCGAAGAAC	138243
	Oy	3356	ATTTTAAAGACTCTTTTAAACAAAAAGCCATGATTAATCTCATATTATTAACATCAATTT	3415
	Db	138242	ATTTTAAAGACTCTTTTAAACAAAAAGCCATGATTAATCTCATATTATTAACATCAATTT	138183
	Oy	3416	ATTTTGTTTTGAACCTGGAGCATGATTTCTATTGTTTATTAATAATTAATGATGT	3469
	Db	138182	ATTTTGTTTTGAACCTGGAGCATGATTTCTATTGTTTATTAATAATTAATGATGT	138129
RESULT 10				
ACCI27071			198104 bp DNA linear HTG 01-AUG-2002	
LOCUS			Homo sapiens clone RP11-200K12, *** SEQUENCING IN PROGRESS ***, 2	
DEFINITION			unordered pieces.	
ACCSSION			ACCI27071.2 GI:21908383	
VERSION			HTG; HTGS_PHASE1; HTGS_ACTIVEFIN.	
KEYWORDS			human.	
SOURCE			Homo sapiens	
ORGANISM			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
REFERENCE			Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS			1 (bases 1 to 198104)	
			Muzny D.M., Adams C., Adio-obuloja B., Ali-oeman F.R., Allen C.,	
			Albrooke S.L., Amaratunge H.C., Aze J.R., Ayale M., Banks T.,	
			Bardack J., Boyle S., Brieva M., Brown E., Brown M., Bryant N.P.,	
			Burckhardt C., Butcher P., Burkett C., Burrell K.L., Byrd N.C.,	
			Carron T.F., Carter M., Cavazos S.R., Chacko J., Chavez D.,	
			Chen G., Chen R., Chen Z., Chowdry I., Christopoulos C.,	
			Cleveland C.D., Cox C., Coyle M.D., Dahome S.R., David R.,	
			Delaney K.R., Delgado O., Denn A.L., Ding Y., Dinh H.H.,	
			Douthwaite K.J., Draper H., Dugan-Rocha S., Durbin K.J.,	
			Fainstet C., Edgar D., Edwards C.C., Elhaj S., Escoto M.,	
			Falls T., Ferraguto D., Flagg N., Ford J., Foster P., Frantz P.,	
			Gabisi J., Gao J., Garcia A., Garner T., Garza N., Gill R.,	
			Gibbs J., Guevara W., Gunaratne P., Hale S., Hamilton K.,	
			Harris C., Harris K., Hart M., Havlak P., Hawes A., Hernandez J.,	
			Hernandez O., Hodgson A., Hogues M., Holloway C., Hollins L.E.,	
			Homs F., Howard S., Huber J., Huliyil S., Hume J., Jackson L.B.,	
			Jacobson B., Jia Y., Johnson R., Jollivet S., Joudah S., Kovar C.,	
			Karlsson E., Kelly S., Khan U., King L., Korvah U., Kovar C.,	
			Kratovic J., Kureishi A., Landry N., Leal B., Lewis L.C., Lewis L.	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3006)
Strausberg, R.

TITLE
JOURNAL

Direct Submission
Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contract: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mdickpax1@stanford.edu
Dickson, M., Schmitz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAC Plate: 48 Row: B Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.

FEATURES
source

Location/Qualifiers
1. 3006
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:33296 IMAGE:5271278"
/tissue_type="Testis"
/clone_lib="NIH MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescript"
4. 2562
/codon_start=1
/product="Similar to p1w1 like homolog 1 (Drosophila)"
/protein_id="AAH31060.1"
/db_xref="GI:21410561"
/translation="MSGRARYKARGIARSPATSEVGRIGQSPLEPSVDLSNNBASNS
GFLGTRISLNDKVGISDAGSTFEMRGVNDKDFMDLSTCRELAVNRKCKTGS
GIPVLTNLPNDPDPDQWQYHVVYIPDLASRRRLIMLSHSELNKKADGA
11FLSQLEKVELSSETQSGRTKMTITKRELPSSVPCQVENVIIKRLKLS
MYOIGNFYSEPEWELPQHKLSLMPGASISVYERKLEFSAVSEYKLVNLEL
MTACORTGISCFTQCEKQILGLIVLRNNRYSDDIDMSGKPTFFOKDGEI
TYDYKQOYDITVSDNQPMVLVSLKRNDSBQALHLEICELIGLIDQATSD
FOLKAVAKERTLSPSGRQQLARLVNIOENTARBELHMGHRSQISLGRIVP
SEKTIIMODHICOPVSAADMSKIDRFCKILNOSNTWMLICSDTEVYASPLNCIR
VAGSMGFNDYPRKIKVQENPAPAFRAIQYVDVQVLMCISLPSNKTYIDSKIL
SDCPVPSQVCLARTLNKQMMSTATKLAOMTCKLGELEAVEILPLKLMVGDV
KDALSKDWNVACVASVNPRTTRMPSCLIQRTMTDVADELKFMGALKMKYRNH
DLPARLIVRAAGDQKTLLEYVPLQSLSSVSSSNTSRLSVIVKCKPRF
TENNRVQNPILGTIVDSATRNEMWDFILSQVACRTVSPYVYIVNGKPKPDH
MORTFLCHLYNMPGIVSVPAPCOYAKRLFTVAOSIHKEPSELNHLFYL"

CDS

BASE COUNT
ORIGIN

Query Match 19.6%; Score 679.2; DB 9; Length 3006;
Best Local Similarity 56.6%; Pred. No. 4.7e-148;
Matches 1320; Conservative 0; Mismatches 1005; Indels 9; Gaps 3;

419 GATTTCATGATCTGGTGAATACAGAGCAACCTGACCACTGTTAAAGATCAAA 478
Db GACTTATGAGATTGATCTGATCCAGAGAAAATTGGACATGGAATAATGTAA 297
479 ACAGGTTCTTCAGGCACTTATAGTAGTAAAGCTTACCACTTCGCGTGAACCCGT 538
Db ACAGGTTCTTCAGGCACTTATAGTAGTAAAGCTTACCACTTCGCGTGAACCCGT 357
298 ACAGGTTCTTCAGGCACTTATAGTAGTAAAGCTTACCACTTCGCGTGAACCCGT 538
539 CCCAGGAGGCTTATATGATGATGATGATGATGATGATGATGATGATGATGATG 598
Db CAAGCTGGCAGCTATACGATGATGATGATGATGATGATGATGATGATGATGATG 417

599 CTCGCTTCACTCTCTCTTTTCAACAGAGATCTAATGAAAGTCACTCTTTGAT 658
418 CTGAGATATGCTTTACTTTATATGATGATGATGATGATGATGATGATGATGAT 477
659 GGAAGAT 718
478 GGGGCACTCTTTTCTGTCACAAAAGCTAGAAAGAAAGTCAAGATGATCAAG 537
719 ACCCGAATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 778
538 ACTCAAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 597
779 TCACCACTGTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 838
598 TCTCCGCTGATCAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 657
839 TTGCAACAATTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 898
658 ATGACCAATTTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 717
899 AGTTGATGATTTGAGCTGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 958
718 AATATATCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 777
959 CTCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1018
778 TTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837
1019 AACTTTATCATCAGACAGAGAAATATATATATATATATATATATATATATAT 1078
838 GCTCTGTCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 897
1079 GGTATGATGATTTTCAACAGATGATGATGATGATGATGATGATGATGATGATG 1138
898 GGGCTATGCTCTTCAACAGATGATGATGATGATGATGATGATGATGATGATG 957
1139 GACCAAGATCCAGAGACCTTTAAGAAAGCGGCTTGGGCTCACTTCTTGA 1198
958 TCAGTGAAGCCCAACACACTTTGAGAGCGGATGACCGAGATCATATGATGAT 1017
1199 TACTACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1258
1018 TACTACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1077
1259 CAGCCAAAGAGAGCGGCGCTGAGGAGCACTGCGAGGCTGCGATGCTATCT 1318
1078 CTGTTAAG--AAGAGAGAAATGACAAAGTGAAGCTCAGCTGATGATGAT 1134
1319 GAGCTGCTATCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATG 1378
1135 GAGCTGCTTCTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1194
1379 GACTTACCGCTTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1438
1195 GCTGAGCTGAGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATG 1254
1439 ATTGATTAATCATTAATAAGATGATGATGATGATGATGATGATGATGATGATG 1498
1255 GTGACCAATCATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1314
1499 TTTGATTCAGCTTCTGCTTCTCAAGAGATGATGATGATGATGATGATGATGATG 1558
1315 TTTGAGAGCAGATA--TCTGATGCTGCGAGATGATGATGATGATGATGATGATG 1371
1559 GGTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1618
1372 CAAGACCAATAT--GTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1428
1619 GCACCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1678
1429 TGCAGATTTAATGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1488
1679 TATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1738

Db 1489 GAATATGTTGCGGAGAGCTTTCTGAACTGCTTGAAGAGTGTGAGAGTGTTCATGGATT 1548
 QY 1739 CAAATGAGAAAGCAATATGATTGAAGTGGATGACAGAACTGAAGCTTAAAGAGTC 1798
 Db 1549 AATGTGACTACCCCAATATCAAAAGTACAGAAATCCAGCTGCAATTTGTTAGAGCT 1608
 QY 1799 TTACAGCAAAAGTTCACAGACACCCAGATAGTGTGCTGTCTGTCTGTTCAAGTAATCGG 1858
 Db 1609 ATACAGCAATATGTTGATCTGATGTTGAGCTGAGTGTGATGCTGCTTCTAATCAG 1668
 QY 1859 AAGGCAAAATACGATCTATTAAAAAATACCTGTGTACAGATTCCTTACCCCAAGTCAG 1918
 Db 1669 AAGACCTATTATGATTTCCATTAATAAATTTTGAAGCTGAGCTGCGGAGTCCCAAGCCAA 1728
 QY 1919 TGTGTGTCGCGGCACTTAGGCAACAGCAAACTGTGTCATGCGCAATGCTCAAAAGATT 1778
 Db 1729 TGTGTGTCGTCGCGGCACTTAGGCAACAGCAAACTGTGTCATGCGCAATGCTCAAAAGATT 1778
 QY 1979 GCCCTACAGATGAATGCAAGATGGAGGAGAGCTCTGGAGGTGACATTCGCCCTGGAAG 2038
 Db 1789 GCTATGCAAGTACTTGCACCTCGAGGCGAGCTGTGGGCTGTGAAATACCTTTAAAG 1848
 QY 2039 CTCGTGATGCTGTCGATCGATTGTTACCATGACATGACAGCTGGGCGGAGTCAATC 2098
 Db 1849 TCCCTGATGTCGTCGCTGATGCTGTTAAAGATGCACTCAGCAAGGACCTGATGCTT 1908
 QY 2099 CGAGATTTGTCGACGATCAATGAAGGATGACCGGCTGTTCTCAGCTGATATTT 2158
 Db 1909 GTTGGATGTCGTCGAGTGTAACTGTTAACTGTTAACTGTTAACTGTTAACTGTTAACT 1968
 QY 2159 CAGCATAGAGGACAGAGCTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2218
 Db 1969 CAGAGACATGATGATGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2218
 QY 2219 GTTGGATAGCTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2278
 Db 2029 AATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2278
 QY 2279 GGAGAGCGGACGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2338
 Db 2089 GGGGATGTCGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2338
 QY 2339 AATTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2398
 Db 2149 CGAGATTCAGCTCAAAATACAGCTCAAGCTGTCGCTGATGTCGCTGATGTCGCTG 2398
 QY 2399 AATCCAGATTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2458
 Db 2209 ATGCCAGATTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2458
 QY 2459 ATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2518
 Db 2269 GTGGATTCAGAGCAACAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2518
 QY 2519 AGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2578
 Db 2329 TGCGGGGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2578
 QY 2579 CCAGACACATACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2638
 Db 2389 CCGACCATATGACAGAGCTTACATTAATTTGTCGCTGATGATGATGATGATGATGATG 2638
 QY 2639 GTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2698
 Db 2449 ATAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2698
 QY 2699 AGTATTCAGAGAGCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2752
 Db 2509 AGCAATTCATAAAGAACCCAGTCTGGAATAGCAACCACTCTCTCTACCTGCTGA 2562

RESULT 12

AK093475
 LOCUS
 DEFINITION
 ACCSSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 source
 CDS
 BASE COUNT
 ORIGIN

AK093475 3138 bp mRNA linear PRI 15-JUL-2002
 Homo sapiens cDNA FLJ36156 fis, clone TESTI2025546, weakly similar
 to Mus musculus mRNA for MIWI (piwi).
 AK093475
 AK093475.1 GI:21752360
 oligo capping; fis (full insert sequence).
 Homo sapiens testis cDNA to mRNA, clone lib:TESTI2
 clone:TESTI2025546.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,
 Ono,Y., Hotuta,T., Hiraoka,S., Murakawa,K., Takiguchi,S.,
 Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,
 Yamashita,H., Chiba,Y., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
 Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y.,
 Saito,K., Nishikawa,T., Kimura,K., Matsuo,K., Nakamura,Y.,
 Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Takahashi-Fujii,A.,
 Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
 Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 3138)
 Isogai,T. and Yamamoto,J.
 Direct Submission
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.
 Location/Qualifiers
 1. 3138
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="TESTI2025546"
 /tissue type="testis"
 /clone lib="TESTI2"
 /note="cloning vector: pME18SFL3"
 152..2710
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAC04179.1"
 /db_xref="GI:21752361"
 /translation="MSGRAVRKARGIARSPATVGRIOASPLPRVDLSNNEASSNS
 GFLGTSTRISTNDKYGISGSDGATPMERGVKNQKQDFMDLSICTREKLAHVNRCKTGS
 GIPVLVTLNLDLPQDQMLQVHVTVIPLASRRRLIALYSHLSNEAKAFDGA
 ILPLSKLEEKVTELSSTQGETIKMTITLRELPSPPVCIQVFNIIIPRKILKLS
 MYIGNFVNPSEPEIPIQHLKSLWPGFSAISYFERKLLFSADVSVKVLRNTEVLEF
 MYALCGTGLSCTCTCKQLIGLIVLTRYNNRTYSIDIDWSVKPHTFKEDGTEI
 TVDYQKQYDITVSDNLQPMVLSLKKRDNSEAOHLIPELFCGLTQTQATSD
 FQMKVAEKTLSPSGRQRLAVDNIORNTNARFELETWGLHSGSLGSLUTGRVLP
 SKILMQDHIQCPVSAADSKDITCKILNAQSLNWLILCSKRTSYVAESFLNCLR
 VAGSGFVNDYPKIIVKQENPAFVRAIQVYVDPVQVLCILPSNOKTYTYSIKYL
 SDCDPSVQCVLARTLNKQMSIATKIAMQMTCKLGGELWAVEIPLKSLMVVGDV
 CKDALSKDVMVGVGVASVNPRTIWFSCILQRTMTDADCLKVFMTGALKWKYKYNH
 DLPARIIVYRAGVGGQLETLIEYEPQLLSVAESGNSSTSRSLSVIVVRKCMRPF
 TEMPTVONPLGTVDSEATRNEDWDFYILISQVACRTVSPPTVYVNVVDNGLKPDH
 MORLTFKLCHLYINWPGVSPVAPCOYAHKLTELVAQSIHKPSLSLANELHYL"

Query Match 19.6%; Score 679.2; DB 9; Length 3138;
 Best Local Similarity 56.6%; Pred. No. 4.8e-148;
 Matches 1320; Conservative 0; Mismatches 1005; Indels 9; Gaps 3;

419 GATTTCATGATCTGTTGATGATCAAGGCAAGCACTGATGATTAAGATCAAAA 478
 386 GACTTTATGATTTGATGATCTGATCAAGAGAAATATGGCAATGTAGAAATTTGAAA 445
 479 ACAGTTCTTCAGGACTTATAGTAAAGTTAAAGCACTTAACATTTCCGCTGACATCCGT 538
 446 ACAGTTTCAAGTGAATACCTGTGAACTGTGTAACAACTCTTAACTTAAGATTTTCC 505
 539 CCCCAGTGGGCTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 598
 506 CAAGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 565
 599 CTCCGTTAGCTCTTCTTTTCAACGAGATCTAATTTGAAGTGTCAATGCTTTGAT 658
 566 CTGAGATGCTTTTACTTATAGTCAATGATGATGATGATGATGATGATGATGATGAT 625
 659 GGAACGATATATTTTACTTAAAGCACTAAGCAAGGATGATGATGATGATGATGATGAT 718
 626 GGTGCTCCTTTTCTGTCACAAAGCTGAGAAAGGATGATGATGATGATGATGATGAT 685
 719 ACCGGAATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 778
 686 ACTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 745
 779 TCACCAATCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 838
 746 TCTCCCTGTCATCCAGGCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 805
 839 TTGCACAATTTGAGCAAGATTTATTTAACTCAATGATGATGATGATGATGATGATGAT 898
 806 ATGTACCAATTTGAGCAAGATTTATTTAACTCAATGATGATGATGATGATGATGATGAT 865
 899 AGGTGATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 958
 866 AATATATCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 925
 959 CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1018
 926 TTTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 985
 1019 AACTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1078
 986 GCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1045
 1079 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1138
 1046 GGGCTCATTTGCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1105
 1139 GACGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1198
 1106 TCAAGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1165
 1199 TACTACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1258
 1166 TACTACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1225
 1259 GAGCCCAAG 1318
 1226 CTGTTAAAG--AAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1282
 1319 GAGCTGCTGATCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1378
 1283 GAGCTGCTGCTTCTAAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1342
 1379 GACTTACGCTGCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1438
 1343 GCTGTGCTGAG 1402
 1439 ATTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1498
 1403 GTGAGCAATCAAG 1462
 1499 TTTGATTCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1558

1463 TTTGGAACCGAGAT--TCTCTGATGCTGCGGATGCTGCTTCAAGAAAAATTAATG 1519
 1559 GGTGAAAAAATCTTTGATTAATATCAATATTTGCAATGTTGTCAAAGAAACAAGAGT 1618
 1520 CAAGACCAATAT--GTCAACCTGTGTCTGTCTGATGATGATGATGATGATGATGATGAT 1576
 1619 GACACATTAATTAATGATTAAGCACTGATGATGATGATGATGATGATGATGATGATGAT 1678
 1577 TCAAGATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1636
 1679 TATGAGAGAGCAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1738
 1637 GATATGTTGCGGAGAGCTTCTGAACTGTTGAGAGATGATGATGATGATGATGATGATGAT 1696
 1739 CAATGAGAAAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1798
 1697 AATGTGACTACCCCAAAATCAATTAAGATGATGATGATGATGATGATGATGATGATGATGAT 1756
 1799 TTAACAGAAAAAGTCAAG 1858
 1757 ATACAGCAATATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1816
 1859 AAGACCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1918
 1817 AAGACCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1876
 1919 TGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1978
 1877 TGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1936
 1979 GGCCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2038
 1937 GCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1996
 2039 CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2098
 1997 TCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2056
 2099 GCAAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2158
 2057 GTTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2116
 2159 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2218
 2117 CAGAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2176
 2219 GCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2278
 2177 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2236
 2279 GAGAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2338
 2237 GGGAT 2296
 2339 AATTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2398
 2297 GCAAGATCGGCTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2356
 2399 AATTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2458
 2357 ATGCAAGATTTCTTAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2416
 2459 ATTGAT 2518
 2417 GTGATTCAG 2476
 2519 AAGAT 2578
 2477 TGCCTGAG 2536
 2579 CCAAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2638

1677 ATTATGAGCAGCAATTCATTGATACAAAATCTATTTAAAGTTACCAAGCCATGGGCA 1736
 953 CTGATATATCTCCGAGAGCTTCTGAACTGCTTGAGAAAGTTACAGGTTCCATGGGAT 1012
 1737 TGCATAATGAGAAAGAAATATGATGTAAGTGAACAGACGAAAGCCACTGATTAAG 1796
 1013 TTAATGTGAGCTACCCCAAAATCATTAAGTAAGAAATTCAGCTGCATTTTGTAGAG 1072
 1797 TCTTACAGAAAGGTCACAGAGACCCAGATAGTTGTCTGTCTGTCTCAAGTAATC 1856
 1073 CTATACAGCAATATGTTGATCTGATGTTGATGATGATGATGATGATGATGATGATG 1132
 1857 GGAAGCAAAATACAGTGTATTTAAATAATCTGTGAACGATTTGCCCTTACCCCAAGTC 1916
 1133 AGAAGACCTATATGATTCATTTAAATAATTTAGGCTCAGACTGCCAGCTCCCAAGCC 1192
 1917 AGTGTGTGTGGCCGACCTTAGGCAACAGCAAACTGTACCTGCTCAAAAGA 1976
 1193 AATGT 1252
 1977 TTGCCCCACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2036
 1253 TCGCTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1312
 2037 AGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2096
 1313 AGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1372
 2097 TCCGAGATTTTGTGACAGATCAATGAAAGGATGACCCGCTGTTCTCAAGCTGATAT 2156
 1373 TTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1432
 2157 TTGAGATGAGAGACAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2216
 1433 TTGAGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1492
 2217 GGGCTTGAATAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2276
 1493 ACAATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1552
 2277 TAGGAGACGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2336
 1553 TAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1612
 2337 TAAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2396
 1613 TGGCAGATCCAGCTCAAAATCCAGCTCAAGACTGTGATGATGATGATGATGATGAT 1672
 2397 TGAACACAGATTTTGTGCTCAATCTGAGAGAGACTTCAAGAAATCCACTTCTGAAAC 2456
 1673 GCATGCCAGATTTCTTACCGAAATGAACCGCACTGACAGAAACCCCTTGGCACTG 1732
 2457 TTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2516
 1733 TTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1792
 2517 TGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2576
 1793 CCGTCCGGGAACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1852
 2577 AGCCAGACCACTACAGAGCTTGAACCTTCAAGCTGTGCAATCTTATTAACAATGAGC 2636
 1853 AGCCAGACCACTACAGAGCTTGAACCTTCAAGCTGTGCAATCTTATTAACAATGAGC 1912
 2637 GTGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2696
 1913 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1972
 2697 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2752
 1973 AAGCATTAATAAGAACCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2028

RESULT 14

AY014899

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

AY014899 3057 bp mRNA linear INV 02-JAN-2001

Strongylocentrotus purpuratus seawater, complete cds.

AY014899

AY014899.1 GI:12007639

Strongylocentrotus purpuratus.

Strongylocentrotus purpuratus.

Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;

Echinoidea; Echinoidea; Echinacea; Echinoida;

Strongylocentrotidae; Strongylocentrotus.

1 (bases 1 to 3057)

Rodriguez, A.J. and Bonder, E.M.

Seawi - Cloning and Molecular Characterization of a Sea Urchin

Homolog of Pw1

Unpublished

2 (bases 1 to 3057)

Rodriguez, A.J. and Bonder, E.M.

Submitted (28-NOV-2000) Biological Sciences, Rutgers

University-Newark, 101 Warren Street Room 135, Newark, NJ 07102,

USA

Location/Qualifiers

1..3057

/organism="Strongylocentrotus purpuratus"

/db_xref="taxon:7668"

181..2745

/product="seawi"

/codon_start=1

/db_xref="GI:12007640"

/translation="MDRRPGEOAPPPYQQAAPROEASGMPVDRGSRGTTOPPG

AVPKQEPREGLTSQIAKAIAGSGEAGVAVDORRRRDEIELEPTTRENVYKAI

AGDKIALFANGFKLTKRDMQYQYRDFEELINPARRALLGSHALFKGLTDM

DMVSLFLKAKVNLSEAKDGSNNVLCYHVAATLPMAPNTLHLNVLFRCLMI

NMEQVRRVYPTAIDIKHGLQIMGFVTSIIQYEDVLLSDLSHKYLRQTVS

VWMDLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VYDYLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VWMDLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VYDYLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VWMDLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VYDYLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VWMDLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VYDYLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VWMDLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VYDYLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VWMDLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VYDYLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VWMDLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VYDYLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VWMDLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VYDYLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VWMDLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VYDYLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VWMDLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VYDYLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

VWMDLFRNRRGFRPEITKMLIGQIVLTKNNKTYRVDIDPDTPADTETSGPVS

AUTHORS Pyle,R.A., Xu,J. and Secrist,H.
TITLE Compositions and methods for the therapy and diagnosis of colon
cancer

JOURNAL Patent: WO 0212280-A 634 14-FEB-2002;
CORIXA CORPORATION (US)

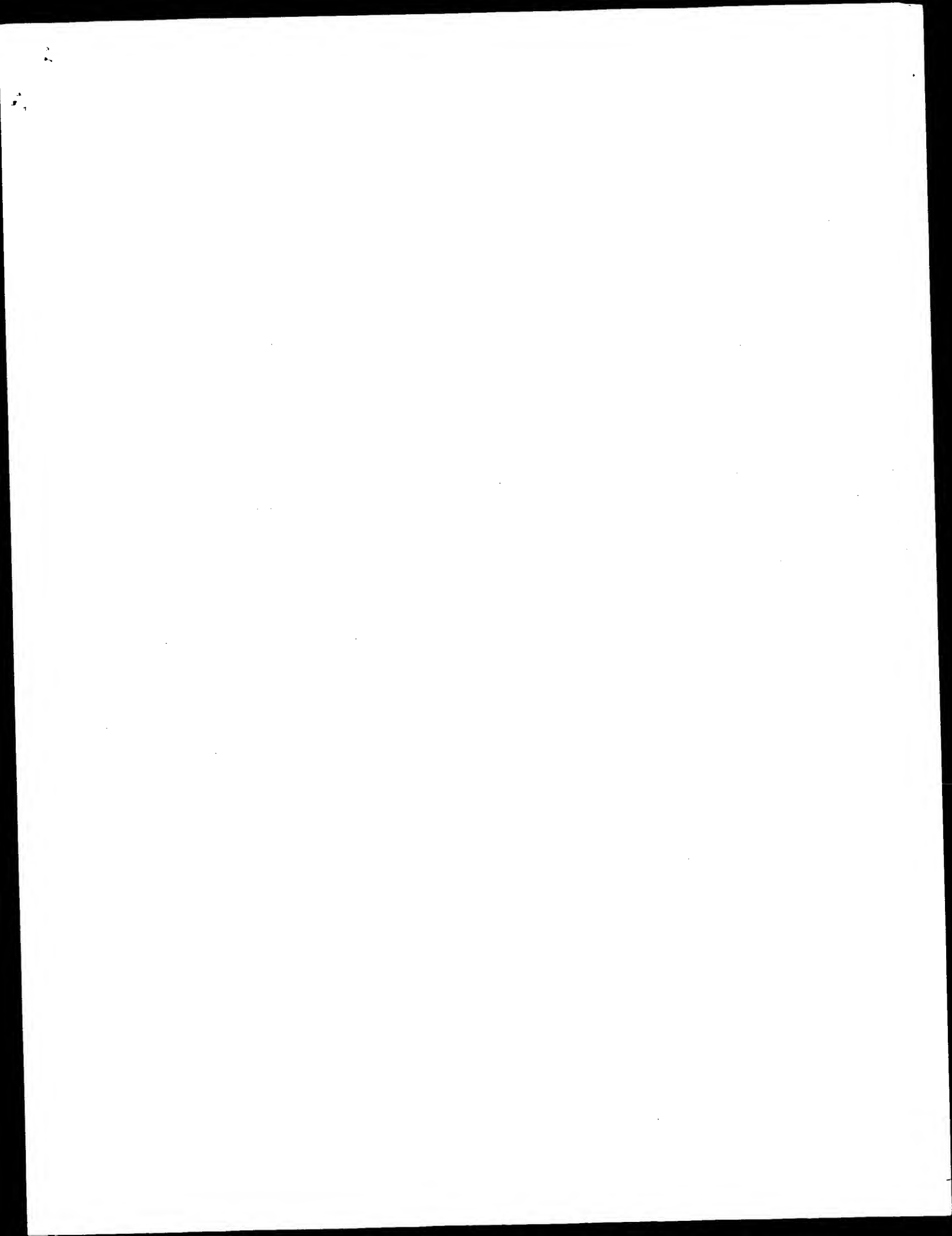
FEATURES Location/Qualifiers
source 1..590

BASE COUNT 209 a 105 c 65 g 208 t 3 others
ORIGIN /organism="Homo sapiens"
/db_xref="taxon:9606"

Query Match 15.9%; Score 550.4; DB 6; Length 590;
Best Local Similarity 99.3%; Pred. No. 5.7e-118;
Matches 551; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2915 TTATTTTATATGTAATAAATTAGATTATTTATTTATCTCTGTTGTTCTCATAGATATTT 2974
DB 590 TTATTTTATATGTAATAAATTAGATTATTTATTTATCTCTGTTGTTCTCANAGANATTT 531
QY 2975 TGTGAGCATTTTGTGTTTATTTTGAAGAAATGTGATAGATATTGTTAGTATAAAC 3034
DB 530 TGTGAGCATTTTGTGTTTATTTTGAAGAAATGTGATAGATATTGTTAGTATAAAC 471
QY 3035 AGACTCTCGAGAGATTTGAAATGTGTTGAGATTTACTTAAAGTACTTTGAGAGT 3094
DB 470 AGACTCTCGAGAGATTTGAAATGTGTTGAGATTTACTTAAAGTACTTTGAGAGT 411
QY 3095 GAGCAAGTCTTATTAATTAATTAATTTTGAATACCTGTTGAAATTT 3154
DB 410 GAGCAAGTCTTATTAATTAATTAATTTTGAATACCTGTTGAAATTT 351
QY 3155 AAAGAGATAGAGGCGTAAGTAGATGCTCACTACACCATAGTGGGTTTCAGCTC 3214
DB 350 AAAGAGATAGAGGCGTAAGTAGATGCTCACTACACCATAGTGGGTTTCAGCTC 291
QY 3215 ATATCTTAAGATTAAGAGTACTATTAATTAATTAATTAATTAATTAATTAATTAAT 3274
DB 290 ATATCTTAAGATTAAGAGTACTATTAATTAATTAATTAATTAATTAATTAATTAAT 231
QY 3275 GCTGATTTTATTTGCGAGGGGCTAGGTTGTATGGAGTAAAAAACAATGAAAT 3334
DB 230 GCTGATTTTATTTGCGAGGGGCTAGGTTGTATGGAGTAAAAAACAATGAAAT 171
QY 3335 TTTTAAATGTCCAAAGAAACATTTTAAGACTCTTAACAAAAAGGCCATGATAATC 3394
DB 170 TTTTAAATGTCCAAAGAAACATTTTAAGACTCTTAACAAAAAGGCCATGATAATC 111
QY 3395 TCTATATTAACATCATTTATTTTGTGAACTGGACATGATTCATTTGTTATA 3454
DB 110 TCTATATTAACATCATTTATTTTGTGAACTGGACATGATTCATTTGTTATA 51
QY 3455 AAATTAATTTGATGT 3469
DB 50 AAATTAATTTGATGT 36

Search completed: January 9, 2003, 20:54:00
Job time : 6591 secs



XX

PA (UYDU-) UNIV DUKE.
 XX Lin H;
 XX WPI: 2000-412085/35.
 DR P-PSDB; AAY90235.
 XX
 PT Piwi family nucleic acids, polypeptides, and antibodies, useful in gene
 therapy of diseases such as cancer and in various research and
 diagnostic applications -
 XX
 XX Claim 19; Page 189-194; 201pp; English.
 XX
 CC This sequence encodes the human piwi family protein, designated
 hiwi. The piwi family nucleic acids and polypeptides are used in gene
 therapy of diseases such as cancer and also in various research and
 CC diagnostic applications. The sequences can also be used to treat
 CC tissue dystrophy, anaemia, immunodeficiency, and male infertility.
 XX
 SQ Sequence 3472 BP; 1044 A; 694 C; 796 G; 935 T; 3 other;

Query Match 99.9%; Score 3469; DB 21; Length 3472;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGGAATGGGCGCCGACGCTGCTCCCGCCGCCATGGCGCGCGGGAATTCGAT 60
 DB 1 GGGGGAATGGGCGCCGACGCTGCTCCCGCCGCCATGGCGCGCGGGAATTCGAT 60

QY 61 TCCATCCTAATACGACTCACTATAGGCTCGAGCGCGCGCGCGGAGTCTGAGTGCA 120
 DB 61 TCCATCCTAATACGACTCACTATAGGCTCGAGCGCGCGCGGAGTCTGAGTGCA 120

QY 121 AGAACGAGGACTAGGCGGAGGCGGCGGTCGAAGAAATAGAAACAAATGACTGGGAGG 180
 DB 121 AGAACGAGGACTAGGCGGAGGCGGCGGTCGAAGAAATAGAAACAAATGACTGGGAGG 180

QY 181 CCAGCCAGAGCCAGAGGAGGCGCGGTCAGGAGAGAGCGGAGTGGTGGGCTCCAC 240
 DB 181 CCAGCCAGAGCCAGAGGAGGCGCGGTCAGGAGAGAGCGGAGTGGTGGGCTCCAC 240

QY 241 TGGCAGTCAGCAACCTGGTGTATATTAGGCTTAGGCTCGAGCGCGCGGAGGAGG 300
 DB 241 TGGCAGTCAGCAACCTGGTGTATATTAGGCTTAGGCTCGAGCGCGGAGGAGG 300

QY 301 ATTATTTGGCGGTGAGCGGAGAGAGGAAACAGCAGAGGAGAAACAGCAAGTCAAGGACT 360
 DB 301 ATTATTTGGCGGTGAGCGGAGAGAGGAAACAGCAGAGGAGAAACAGCAAGTCAAGGACT 360

QY 361 CCAGATATCTGCTGGATTTTCAGGAGTTATCGNTAGCAGAGAGGAGGCTCGTAGAGA 420
 DB 361 CCAGATATCTGCTGGATTTTCAGGAGTTATCGNTAGCAGAGAGGAGGCTCGTAGAGA 420

QY 421 TTTTTCATGATCTTGGTGTGAATACAGGCGAGAACCTAGACCATGTTAAAGAAATCAAAAC 480
 DB 421 TTTTTCATGATCTTGGTGTGAATACAGGCGAGAACCTAGACCATGTTAAAGAAATCAAAAC 480

QY 481 AGGTTCTTCAGGCAATATAGTAAGGTTAAGCACTAACCAATTTCCGCTGACATCCCGTCC 540
 DB 481 AGGTTCTTCAGGCAATATAGTAAGGTTAAGCACTAACCAATTTCCGCTGACATCCCGTCC 540

QY 541 CCAGTGGGCTTATATCAGTATCAGATTCATATACCCACTGATGGAGCCAGAGACT 600
 DB 541 CCAGTGGGCTTATATCAGTATCAGATTCATATACCCACTGATGGAGCCAGAGACT 600

QY 601 CCCTTCAGCTCTCTTTTCAACAGCAAGATCTAATTTGGAAGTGTCACTGCTTTTGTGG 660
 DB 601 CCCTTCAGCTCTCTTTTCAACAGCAAGATCTAATTTGGAAGTGTCACTGCTTTTGTGG 660

QY 661 AACGATATTTTACCTTAAAGACTACAGCAAGGTTTCTGAAGTTTCTAGTAAGAC 720
 DB 661 AACGATATTTTACCTTAAAGACTACAGCAAGGTTTCTGAAGTTTCTAGTAAGAC 720

QY 721 CCGGAATGGAGGAGTGTAGGATAACGATCACTTTTAAACAATGAACCTTCCACCTACATC 780
 DB 721 CCGGAATGGAGGAGTGTAGGATAACGATCACTTTTAAACAATGAACCTTCCACCTACATC 780

QY 781 ACCAACTTGTGTGAGTCTTATATATTTTTCAGGAGGCTTTTGAATAATCATGAATTT 840
 DB 781 ACCAACTTGTGTGAGTCTTATATATTTTTCAGGAGGCTTTTGAATAATCATGAATTT 840

QY 841 GCAACAATTTGGAGCAAAATTTATTAACCAATGACCAATTTGATTTTCAAGTCAAG 900
 DB 841 GCAACAATTTGGAGCAAAATTTATTAACCAATGACCAATTTGATTTTCAAGTCAAG 900

QY 901 GTTGTGATTTGGCTCTGCTTCACTACTTCCATCTTCACTATGAAACAGCATCATGCT 960
 DB 901 GTTGTGATTTGGCTCTGCTTCACTACTTCCATCTTCACTATGAAACAGCATCATGCT 960

QY 961 CTGCACTGACCTTTAGCCCAATTAAGTCTTCCGAAGTGAAGTCTTTTGGATTTCAATGTTCAA 1020
 DB 961 CTGCACTGACCTTTAGCCCAATTAAGTCTTCCGAAGTGAAGTCTTTTGGATTTCAATGTTCAA 1020

QY 1021 CTTTATCATCAGACAGAAACATAAATTTTCAAGAAACAAGTTTCCAAAGAAATAATAGG 1080
 DB 1021 CTTTATCATCAGACAGAAACATAAATTTTCAAGAAACAAGTTTCCAAAGAAATAATAGG 1080

QY 1081 TTTAGTGTCTTACCAAGTATTAACATTAAGACATACAGAGTGGATGATTTGACTGGGA 1140
 DB 1081 TTTAGTGTCTTACCAAGTATTAACATTAAGACATACAGAGTGGATGATTTGACTGGGA 1140

QY 1141 CCAGATCCCAAGAGCACCTTTTAAAGAAAGCCGCTCTGGGGTCACTTCTTAGAATA 1200
 DB 1141 CCAGATCCCAAGAGCACCTTTTAAAGAAAGCCGCTCTGGGGTCACTTCTTAGAATA 1200

QY 1201 CTAAGGAGCAATTAACCAAGAGATCAACGAGTCTTGAAGAGCCTGTCTTGTGACGCA 1260
 DB 1201 CTAAGGAGCAATTAACCAAGAGATCAACGAGTCTTGAAGAGCCTGTCTTGTGACGCA 1260

QY 1261 GCCAAGAGAGGCGGCGCTTGGGGGACACTGCCAGGCGCTGCCATGCTCAATCTCTGA 1320
 DB 1261 GCCAAGAGAGGCGGCGCTTGGGGGACACTGCCAGGCGCTGCCATGCTCAATCTCTGA 1320

QY 1321 GCTCTGCTTCTTACAGGTCTAATGATAAATGCGTAATGATTTTAAACGTGATGAAGA 1380
 DB 1321 GCTCTGCTTCTTACAGGTCTAATGATAAATGCGTAATGATTTTAAACGTGATGAAGA 1380

QY 1381 CTTAGCGTTCATACAGACTTAATCCAGAGCAAGAGCGCTGAAAGTGGGAGCATCAT 1440
 DB 1381 CTTAGCGTTCATACAGACTTAATCCAGAGCAAGAGCGCTGAAAGTGGGAGCATCAT 1440

QY 1441 TGATTAATTTCAATAAAGCAATGTTTCAAGGGAGCTTCCAGAGCTGGGGTTGAGCTT 1500
 DB 1441 TGATTAATTTCAATAAAGCAATGTTTCAAGGGAGCTTCCAGAGCTGGGGTTGAGCTT 1500

QY 1501 TGATTTCAACTTACTGCTTCTCAGGAAGATTTTGCACAAAGAAAGATTCCACCAAGG 1560
 DB 1501 TGATTTCAACTTACTGCTTCTCAGGAAGATTTTGCACAAAGAAAGATTCCACCAAGG 1560

QY 1561 TGGAAAAACATTTGATTTACCAATCCCAATTTGAGATTTGGTCCAAAGAAACAGAGGTGC 1620
 DB 1561 TGGAAAAACATTTGATTTACCAATCCCAATTTGAGATTTGGTCCAAAGAAACAGAGGTGC 1620

QY 1621 ACCATTAATTTAGTGTAAAGCCACTAGATTAACCTGCTGTTGATCTATACCGAAGAAATTA 1680
 DB 1621 ACCATTAATTTAGTGTAAAGCCACTAGATTAACCTGCTGTTGATCTATACCGAAGAAATTA 1680

QY 1681 TGAAGCAGCAATTTCAATTTGATACAAAATCTATTAAAGTTTACACCGCATGGCATGCA 1740
 DB 1681 TGAAGCAGCAATTTCAATTTGATACAAAATCTATTAAAGTTTACACCGCATGGCATGCA 1740

QY 1741 AATGAGAAAAAGCAATTAATGATTTGAAGTGAAGAGCACTGAAGGCTTCTTAAGAGCTTT 1800
 DB 1741 AATGAGAAAAAGCAATTAATGATTTGAAGTGAAGAGCACTGAAGGCTTCTTAAGAGCTTT 1800

QY 1801 ACAGCAAAAGGTTCACAGCAGACCCAGATAGTTGCTGCTGCTTCTCAAGTAATCGGAA 1860

PA (UYDU-) UNIV DUKE.

XX Lin H;

DR WPI; 2000-412085/35.

DR P-PSDB; AAY90234.

XX Piwi family nucleic acids, polypeptides, and antibodies, useful in gene
PT therapy of diseases such as cancer and in various research and
PT diagnostic applications.

XX Claim 19; Page 180-185; 201pp; English.

XX This sequence encodes the mouse piwi family protein, designated
CC miwi. The piwi family nucleic acids and polypeptides are used in gene
CC therapy of diseases such as cancer and also in various research and
CC diagnostic applications. The sequences can also be used to treat
CC tissue dystrophy, anaemia, immunodeficiency, and male infertility.

XX Sequence 4064 BP; 1114 A; 978 C; 1077 G; 890 T; 5 other;

Query March 52.8%; Score 1833.6; DB 21; Length 4064;

Best Local Similarity 81.8%; Pred. No. 0;

Matches 2125; Conservative 0; Mismatches 471; Indels 3; Gaps 1;

QY 164 ACAATGACTGGAGAGCCGAGCCAGAGCCAGAGAGGCGCGGTTCAGGAGACAGCG 223

DB 188 AATATGACTGCCAGAGCCGAGCTCGGGCCGCGCGAGGCAAGAGTCAGGAGCGGTG 247

QY 224 CAGCTGGTGGCTCCACTGCCAGTCAGCAACCTGGTTATATTCAGGCTAGGCTCAGCGG 283

DB 248 CAGCATGTGGGGCTGCTCGAGCCAGCAACCTGGGTACATCCACCGAGACCTCAACAG 307

QY 284 CCACGACGAGGGGGAATATTGTCGCGTGGAGCGGAGGAGGAGAA---CAGCAGGAGGA 340

DB 308 TCCCCCAGAGGGGGGACTTGGTTGGCCGAGGACGACAGAGGGGGGATGGTAGTCGAGGCC 367

QY 341 ACAGCCAAAGTCAAGAGACTCCAGATATCTGCTGGATTTTCAGGAGTTATCGNTAGCAGAG 400

DB 368 ACATCCAAGTCACAAAGTCTCAGATCTCAGCTGGTTTCAGGAGCTGCTAGTGGCAGAG 427

QY 401 AGAGAGGTGCTGTAGAGATTTTCATGATCTTGGTGTGAATACAGGCGAGAACCTAGAC 460

DB 428 AGAGAGGGGCTGCCGAGACTTCCATGACNTTGGTGTGAACACACAGACAGAACCTTGAC 487

QY 461 CATGTTAAAGAAATCAAAACAGGTTCTTCAGGCAATATAGTAAGTTAAGCACTAACCAT 520

DB 488 CATGTCAAAGAGTCAAGACAGGCTCTCTGGCATCAITTTGAAGCTAGACACCAACAC 547

QY 521 TTCGGCTGACATCCCGTCCCGAGTGGGCTTATATCAGTATCAGATGACATATTAACCCA 580

DB 548 TTCGGCTGACCTCGCGCCACAGTGGGCTTATCAGTATCAGTATCAGTATCAGTATCAG 607

QY 581 CTGATGAGACGAGACATCCGCTTCAGCTCTCTTTTTCACACAGAGATCTAATGGA 640

DB 608 CTGATGAGGCCCCGAGAGCTTCGCTCCGACTGCTCTCCAGCATGAAGCTCATTTGGA 667

QY 641 AAGTGTCTATCTTTGATGGAACGATTTATTTTACCTTAAAGACTACAGCAAGGTT 700

DB 668 AGTGTCTATCTTTGATGGAACGATTTATTTTACCTTAAAGACTACAGCAAGGTT 727

QY 701 ACTGAAGTTTTAGTAAGACCCGGAATGAGAGGATGTGAGGATAACGATCACTTTAACA 760

DB 728 ACAGAAGTATTCAGTCAGACTCGGAATGGGAAACAGTGAGGATCACCATCACTGACC 787

QY 761 AATGAATCTCCACTACATCACCCTTTGTCAGTTCTTATATATTTATTTTTCAGAGG 820

DB 788 AACGAGCTGCCGCCACCTCGGCCACCTGCGAGTTCTATATATCATCTTCAGGAGG 847

QY 821 CTTTTGAAATCATGAATTTGCAACAAATTTGACGAAATTTATTAACCCAAATGACCCA 880

DB 848 CTCTTGAAATCATGAATTTGCAACAAATTTGACGAAATTTATTAACCCAAATGACCCC 907

QY 881 ATTGATATTCAGTCAAGTTCAGAGTTGGTGGCTTGGCTTCACTACTTCCATCTTCAG 940

DB 908 ATTGATATTCAGTCAAGTTCAGAGTTGGTGGCTTGGCTTCACTACTTCCATCTTCAG 967

QY 941 TATGAAGAAGCATCATCTCTGACCTGAGCTTGGTGGCTTGGCTTCACTACTTCCATCTTCAG 1000

DB 968 TATGAAGAAGCATCATCTCTGACCTGAGCTTGGTGGCTTGGCTTCACTACTTCCATCTTCAG 1027

QY 1001 GTTTTGGATTTCACTTCACTTATCATCAGACAGAGTTCAGTTCAGTTCAGTTCAGTTCAG 1060

DB 1028 GTCTTACACTTCACTTATCATCAGACAGAGTTCAGTTCAGTTCAGTTCAGTTCAG 1087

QY 1061 GTTTTCAAGAAGTAAATAGGTTTGGTGGCTTGGCTTCACTACTTCCATCTTCAGTTCAG 1120

DB 1088 GTTGTGAAGAAGTAAATAGGTTTGGTGGCTTGGCTTCACTACTTCCATCTTCAGTTCAG 1147

QY 1121 GTGATCATATTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1180

DB 1148 GTGATCATATTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1207

QY 1181 GGGTCAAGCTTCTTAGAATACAGAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1240

DB 1208 GAGTCAAGCTTCTTAGAATACAGAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1267

QY 1241 CAGCTGTCTTGGTCAAGCTTCAAGAGAGGCGGCGCTTGGGCGGACACTTGCAGGG 1300

DB 1268 CAGCGGTGTCTGGTCAAGCTTCAAGAGAGGCGGCGCTTGGGCGGACACTTGCAGGG 1327

QY 1301 CTTGCCATGCTCACTTCTGAGCTTCTTACAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1360

DB 1328 CAGCTATGATCTTCTGAGCTTCTTACAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1387

QY 1361 GATTTTACGTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1420

DB 1388 GATTTTCAATGTGATGAGGACCTTGGCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1447

QY 1421 CGTGAAGTGGAGGACTCATTTGATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTT 1480

DB 1448 CGGAGGTGGGCGGCTTCACTGATTCATTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1507

QY 1481 CGAGACTGGGGTTCAGCTTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1540

DB 1508 CGAGACTGGGGTTCAGCTTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1567

QY 1541 ACAGAAGATTCACCAAGTGGGAAAAACATTTGATTAACAATTCACCAATTTGAGAGTTGG 1600

DB 1568 TCTGAGAAGATCCAGCGGCGGAAAGACGTTTGAATACCAATTTGAGAGTTGG 1627

QY 1601 TCCAGAAGAACAGAGGTCGACCATTAATTTAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1660

DB 1628 TCCAGAAGAACAGAGGTCGACCATTAATTTAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1687

QY 1661 ATCTATACCGAAGAAATTTGAAGCAGCCAAATTCATTTGATACAAATTCATTTAAAGTT 1720

DB 1688 ATCTATACCGCAGGAATTTGAAGCAGCCAAATTCATTTGATACAAATTCATTTAAAGTT 1747

QY 1721 ACACAGGATTCGAGGATTCGAAATGAGAAAGCAATTAATTTGAGTTCAGTTCAGTTCAGTTCAG 1780

DB 1748 ACTCAGGATTCGAGGATTCGAAATGAGAAAGCAATTAATTTGAGTTCAGTTCAGTTCAGTTCAG 1807

QY 1781 GAAGCTTACTTAAAGTCTTACAGCAAAAGTTCAGCAGACACCCAGATAGTTCGTTGT 1840

DB 1808 GAAGCTTACTTAAAGTCTTACAGCAAAAGTTCAGCAGACACCCAGATAGTTCGTTGT 1867

QY 1841 GTTGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1900

DB 1868 GTTGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1927

QY 1901 TGCCCTTCCCAAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1960

DB 1928 TGCCCTTCCCAAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1987

QY 1961 GCCATTGCTACAAAGATTTGCCCTTACAGATGAACTGCAAGATGGGAGGAGGCTTCGAGG 2020

Db 110 TCTATATTAACATTACTATTATTTTGTGGAACCTGGACATGATCTATTGTTATA 51
 QY 3455 AAATAAAATTGATGT 3469
 Db 50 AAATAAAATTGATGT 36

RESULT 4

ABK44483
 ID ABK44483 standard; cDNA; 501 BP.
 XX
 AC ABK44483;

XX
 DT 05-JUN-2002 (first entry)

XX cDNA encoding colon tumour protein, SEQ ID No 34.

XX Human; colon tumour; vaccine; colon cancer; immunogenic;
 KW immunotherapy; gene; ss.

XX Homo sapiens.

XX WO2000212328-A2.

XX 14-FEB-2002.

XX 31-JUL-2001; 2001WO-US24218.

XX 03-AUG-2000; 2000US-23283P.

XX 28-MAR-2001; 2001US-279763P.

XX 29-JUN-2001; 2001US-302051P.

XX (CORI-) CORIXA CORP.

XX King GE, Meagher MJ, Xu J, Secrist H;

XX WPI; 2002-241739/29.

XX New colon cancer polypeptides and polynucleotides, useful as vaccines,
 PT for diagnosing, preventing, and treating colon cancer, and as markers
 PT for the progression of cancer -

XX Claim 1; SEQ ID No 34; 147pp; English.

XX The invention relates to polynucleotides encoding colon tumour proteins.
 CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
 CC compositions, such as vaccines, for the diagnosis, prevention, and
 CC treatment of colon cancer. Polynucleotide sequences may be used as
 CC hybridisation probes or primers, and in the design and preparation of
 CC ribozyme molecules for inhibiting expression of tumour polypeptides and
 CC proteins in tumour cells. The compositions are useful for stimulating an
 CC immune response against cancer, particularly for the immunotherapy of
 CC colon cancer, and as markers for the progression of cancer.
 CC ABK44450-ABK46237 represent coding sequences of human colon tumour
 CC proteins of the invention.

CC Note: With the exception of SEQ ID No 1 and 2, the sequence data
 CC for this patent did not form part of the printed specification but was
 CC supplied by the European Patent Office.

XX Sequence 501 BP; 134 A; 81 C; 93 G; 193 T; 0 other;

Query Match 13.2%; Score 459.4; DB 24; Length 501;
 Best Local Similarity 99.8%; Pred. No. 5.6e-105;
 Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2695 CCAGAGTATTCACAGAGCCAAATCTGTCAGTCAACACCGCTTTACTACTTAACC 2754
 Db 1 CCAGAGTATTCACAGAGCCAAATCTGTCAGTCAACACCGCTTTACTACTTAACC 60

QY 2755 TGCAGAGAGGATGACCGCGCTTTCTTTTGAATGACTTTGGGATTTTAAAGCTTT 2814
 Db 61 TGCAGAGAGGATGACCGCGCTTTCTTTTGAATGACTTTGGGATTTTAAAGCTTT 120

QY 2815 TATTACTTTTTTAACTGTTATCTTCTCGATGAAACTTTGGGAGGGGATTAGGAGA 2874
 Db 121 TATTACTTTTTTAACTGTTATCTTCTCGATGAAACTTTGGGAGGGGATTAGGAGA 180
 QY 2875 TCTAGCATTTTATCTAGCATTTGCTATTACCGGCTTCCCTTATTATATGTAATAAATT 2934
 Db 181 TCTAGCATTTTATCTAGCATTTGCTATTACCGGCTTCCCTTATTATATGTAATAAATT 240
 QY 2935 AAGATTTTATATTTATCTTCTTCTTCTCATAGATATTTTGTGAGCATTTTGTGTTA 2994
 Db 241 AAGATTTTATATTTATCTTCTTCTTCTCATAGATATTTTGTGAGCATTTTGTGTTA 300
 QY 2995 TTTTGAAGAAATGTGGATAAGATACCTTGGTAGTATAAAACAGACTCTCTGAGAGTATTG 3054
 Db 301 TTTTGAAGAAATGTGGATAAGATACCTTGGTAGTATAAAACAGACTCTCTGAGAGTATTG 360
 QY 3055 AAATGTTTGGAGATTTTACTTAAACGTTACTTTCAGGAGTGAGCAAGTCTCTACTTATAAA 3114
 Db 361 AAATGTTTGGAGATTTTACTTAAACGTTACTTTCAGGAGTGAGCAAGTCTCTACTTATAAA 420
 QY 3115 CCTATATTAACCTTTATTTTGTGAGATACCTGTTTGAATTTA 3155
 Db 421 CCTATATTAACCTTTATTTTGTGAGATACCTGTTTGAATTTA 461

RESULT 5

AAS25525/c

ID AAS25525 standard; cDNA; 460 BP.

XX AC AAS25525;

XX DT 07-NOV-2001 (first entry)

XX Human ovarian PCR-subtracted cDNA library clone #1610.

XX Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;
 KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;
 KW primer; probe.

XX OS Homo sapiens.

XX PN WO200157207-A2.

XX PD 09-AUG-2001.

XX PF 05-FEB-2001; 2001WO-US03733.

XX PR 04-FEB-2000; 2000US-0180403.

XX PR 28-MAR-2000; 2000US-0192745.

XX PA (CORI-) CORIXA CORP.

XX PI Algate PA, Mannion J;

XX WPI; 2001-488875/53.

XX New polynucleotides encoding ovarian tumour proteins, useful for
 PT treating ovarian cancer, and as probes, primers, and markers of cancer
 PT progression -

XX Example 1; page 374; 378pp; English.

XX The invention comprises compositions used for the therapy and diagnosis
 CC of ovarian cancer. The compositions comprise one or more ovarian tumour
 CC proteins, their associated polynucleotides, or immunogenic portions of
 CC the proteins. The ovarian tumour polynucleotides and polypeptides are
 CC useful for stimulating and/or expanding T cells specific for a tumour
 CC protein. They are also useful for inhibiting the development of cancer in
 CC a patient with an ovarian tumour DNA or protein by incubating isolated
 CC T-cells allowing them to proliferate, and administering to the patient.
 CC The sequences can be used as markers for cancer, for example, to monitor
 CC ovarian cancer progression. Probes and primers are useful in nucleic acid
 CC hybridisation, in detecting the presence of complementary sequences in a

XX 18-DEC-2001 (first entry)
 DT cDNA sequence encoding mammalian Sp916.
 XX Mammalian; reproductive-specific protein; male infertility; gene therapy;
 XX spermatogenesis; sperm count disorder; anti infertility;
 KW reproduction; ss.
 XX Mammalia.
 OS WO200166752-A2.
 PN 13-SEP-2001.
 PD 07-MAR-2001; 2001WO-US07371.
 XX 07-MAR-2000; 2000US-0187518.
 PR 12-JAN-2001; 2001US-0261557.
 XX (WHEED) WHITEHEAD INST BIOMEDICAL RES.
 PA Wang PJ, Page DC;
 XX WPI: 2001-570774/64.
 DR P-PSDB; AAU07866.
 XX Novel reproduction-specific protein, useful for treating disorders of
 PT reduced sperm count, enhancing/increasing sperm count and/or sperm
 PT activity -
 XX Claim 1; Fig 15; 151pp; English.
 XX The present invention relates to the isolation of novel mammalian and
 CC human reproductive-specific proteins (AAU07859-AAU07899), and the
 CC nucleic acids encoding them. The nucleic acids encoding
 CC reproductive-specific proteins are useful for diagnosing infertility,
 CC which is a result of reduced sperm count, reduced sperm motility,
 CC malformed sperm or combinations of these. The sequences of the invention
 CC are useful as markers for spermatogenic cells, for identifying genes or
 CC proteins characteristic of male infertility, diagnosing or aiding in
 CC the diagnosis of infertility in men, and for contraception in which
 CC sperm production or sperm count is reduced or defective sperm is
 CC produced. Antibodies to reproductive-specific proteins are useful for
 CC determining the presence of these proteins in a sample obtained from a
 CC man being assessed for infertility, for identifying the expression of
 CC genes in particular cell type or particular developmental stage, for
 CC studies of spermatogenesis, and for immunofluorescence of germ cells or
 CC in Western blots for assessing the presence of the protein the antibody
 CC binds. The sequences of the invention are also useful for treating
 CC disorders of reduced sperm count, and for increasing sperm count and/or
 CC sperm activity. The nucleic acids of the invention are useful in gene
 CC therapy. AAS13623-AAS13647 represent cDNA sequences encoding for the
 CC mammalian reproduction-specific proteins of the present invention.
 XX Sequence 3649 BP; 947 A; 841 C; 911 G; 950 T; 0 other;

Query Match 10.9%; Score 377; DB 22; Length 3649;
 Best Local Similarity 52.5%; Pred. No. 7.1e-84;
 Matches 974; Conservative 0; Mismatches 862; Indels 21; Gaps 6;

QY 914 CTGGGCTTCACTTCCATCCCTTCCAGTATGAAACAGCATCATGCTGCACTACGCTT 973
 DB 1 CTGGGCTATGCTGGGCTAGTATCCGAGGAGCAGACGGGGGCTCTTCTGCTGCTCATGTC 60
 QY 974 AGCCATAAAGTCTTCCAGTGGAGCTGTTTGGATTTCACTTCAACTTTTATCATCAG 1033
 DB 61 TCTCATAGGTGATTCGGACGACTCTGTGCTGGATGTCATGCTATCTACAGAG 120
 QY 1034 ACAGAGAACATAAATTTCCAGAACAGTTTCCAAAGAAATTAATAGTTTGTGTTT 1093
 DB 121 AACAGAGGACAC---TTCCAGGACGATGTCAGCAAGCTTCTGTTGGCAGCATTTGTCATC 177

QY 1094 ACCAAGTATACATTAAGACATACAGAGTGGATGATATTTGCTGGGACCAAGTCCCAAG 1153
 DB 178 ACGGCTTACAAATCGTACCTCCGAATCGATGATGGAGTGGAAACAGACCCCTAAA 237
 QY 1154 AGCACCTTTAAGAAAGCGCGCTCTGGGGTCTAGCTTTCTTAGAATACTACAGGAAGCAA 1213
 DB 238 GACAGCTTTGTGTCATGTCGGAGCGGAAGAAATCACATTTCTGGAATACTACAGCAAAAC 297
 QY 1214 TACAACCAAGAGATCACCGACTTGAAGCAGCTGTCTTTGGTCAAGCCAGCCCA---AGAGA 1270
 DB 298 TATGGGATCAAGTCAAGGAAGATGACACCGCTGTGATCCACCGGCCAGTGAGAGA 357
 QY 1271 AGCGGGGCCCTGGGGGACACTGCCAGGCGCTGCCATGCTCATTTCTGAGCTCTGCTAT 1330
 DB 358 CAGAAATACCATGGCATGTTGCTGAAGGGGAGATCTCTGCTGCTGCCGAGCTCTCTTC 417
 QY 1331 CTTACAGGTCTAACTGATAAAATCGGTAATGATTTTAACTGATGAAGACTTTAGCCGTT 1390
 DB 418 ATGACGGGATCCCTGAGAGATGAAGAAGACTTCAGGGCCATGAAGACTTTGACTCAG 477
 QY 1391 CATACAAGACTAATCCAGAGCAAGGACGCTGAAGTGGAGACTCATTTGATTAAT 1450
 DB 478 CAGATTAACTTCCAGCCCAAGCAGACACCGGTGCTTTGGAATGCTCTGCTGACAGAAAT 537
 QY 1451 CATAAAACGATATGTTCAAGGGAGCTTCGAGACTGGGGTTTGAGCTTTGATTTCAAC 1510
 DB 538 TCACAAAACGAGACAGCCCAATGAGCTGACCCGCTGGGGGCTCAGTCTGCAATAAGAT 597
 QY 1511 TTACTGCTCTTCTCAGGAAGAAATTTTGAACAGAAAAGATTCACCAAGGTGGAAAACA 1570
 DB 598 GTCCACAAGATTGAAGTTCGGCTTCTGCCAATGGAGGATCACTTTAAGAACACTTCA 657
 QY 1571 TTTGATTAATCCAAATTTGCGAGATTGTCACAAAGAAACAGAGGTGACCAATTAAT 1630
 DB 658 TTTGT---CACATCGGAGGCTGAACTGGGTTAAGGAAGTGCACAGAGATGCTTCCAT 714
 QY 1631 AGTGTAAAGCCACTAGATACTGCTGTTGATCTATACGCGAAGAAATTTATGAAGAGCC 1690
 DB 715 CTAATATTTCCATGCAATTTCTGGGCACTCTTTATTCAAAGAGAGCAATGGACCAAGCC 774
 QY 1691 AATTCATTGATCAAAAATCTATTAAAGTTTACACAGCCATGGGCATGCAAAATGAGAAA 1750
 DB 775 AGAAGCTGTTTAACTGTTGGAAGAAATGTCGGGCCCATTTGGCATGGCCACACCC 834
 QY 1751 GCATATATGATTGAAGT---GGATGACAGAACTGAAGCTTCTTAAGAGTCTTTACAGAA 1807
 DB 835 CCAGCTGGGTTGAGCTGAAGGATGACCGAATAGAGACCTATATCAGGACCATTCAGTCC 894
 QY 1808 AAGTCAACAGAC 1861
 DB 895 TTAAGTGGAGTTGAGGGGAGAGATACAAATGGTCTGTTGATCATCATGAGGACACCTGAT 954
 QY 1862 GACAAATACGATGCTATTAAAAAATACCTGTGTACAGATTGCCCTACCCCAAGTCAAGTGT 1921
 DB 955 GATCTCTATGAGGCCATCAAGAGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014
 QY 1922 GTGGTGCCCGAACCTTTAGGCAAAACAGCAAAATGTTGATGGCCATTTGCTACAAAGATTGCC 1981
 DB 1015 ATCAATGTCCGAACCATTTGCTGAGCCACAGGCTTCGGAGCGTGGCTCAGAAAATTTTA 1074
 QY 1982 CTACAGATGAATGCAAGATGAGGAGGAGCTCTGAGGGTGGACATCCCTTGAAGCTC 2041
 DB 1075 CTTCAAGTAACTGTAACTGGGTGGTGGAGTCTGGGGAGTGGATATTCCGTGGAACAA 1134
 QY 2042 GTGATGATCTTGGCATTCGATTTTACCATGATGACAGCTGGGGGAGGTTCAATTCGCA 2101
 DB 1135 CTAATGCTGATTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1194
 QY 2102 GGATTTGTTCCAGCATCAATGAAGGAGTGAACCGCTGTTTCTCAAGCTGCAATTTTCA 2161
 DB 1195 GGCTTCTGGCCAGCATAAATCTCACACTCAACAAATGGTGTACTCGAGGGTGGTGTTCAG 1254
 QY 2162 GATAGAGGACAGGAGCTGGTGTAGTGGCTCAAGTCTGCTGCAAGCGGCTCTGAGGGCT 2221


```

Db 1255 ATGCCACATCAGAGATGTTGACAGCGCTGAGAGCTGCTGGTGGTGTCTTGTAAAG 1314
Qy 2222 TGGAAATAGCTGCAATGATGATACATCCCGAGGATCATCTGTTACCGGATGCGTGA 2281
Db 1315 TATTATGAGTGAACCATGCTCTCCAGAGAAATTTGGTGTACCGAGATGAGTGTCT 1374
Qy 2282 GACGGCGAGCTGAAACATCTGTTGAACTACGAAAGGCCAGTTTGGATTTGTAA 2341
Db 1375 GATGGCGAGCTGAAAGAGCTTCCATCTACGAGATCCCTGAGCTGCAAGATGTTTGA 1434
Qy 2342 TCCATTGTTGAGGATTACAACTTAGATTAACGTTATTTGTTGTAAGAAAGAGTGA 2401
Db 1435 GCTTTGATA---ACTACACACCCCAAGATGATGTTGTTGTTGTTGTTGTTGTTGTT 1491
Qy 2402 ACCAATTTTTTGTCTGATCTGTTGAGAGAGATCTTCAAGATCCATCTCTGGAACAT 2461
Db 1492 ACCAATCTGTACCTGCTGCTCTCTGATCTTCTGTAACCCCTCCCGGAGCTGTGTT 1551
Qy 2462 GATGTAGAGTTTACAGACACCAAGATGATGATGATGATGATGATGATGATGATGAT 1521
Db 1552 GATCATACATACATACATGCTGATGATGATGATGATGATGATGATGATGATGATG 1611
Qy 2522 AGTGTAGTGTGTTCTCCACACATTAATGATGATGATGATGATGATGATGATGATG 2581
Db 1612 CAGGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1671
Qy 2582 GACCATATACAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2641
Db 1672 GATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1731
Qy 2642 ATTCGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2701
Db 1732 ATCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1791
Qy 2702 ATTCAGAGAGCGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2758
Db 1792 TTGCATCATGAGCGACCATTCAGAGCTGTGTGGAAACCTGCTGCTGCTGCTGCTG 1848

```

RESULT 8
AAH15960
ID AAH15960 standard; cDNA; 2272 BP.

AAH15960;

26-JUN-2001 (first entry)

Human cDNA sequence SEQ ID NO:14558.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-0116126.

29-JUL-1999; 99JP-0248036.

27-AUG-1999; 99JP-0300253.

11-JAN-2000; 2000JP-0118776.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

(HELI-) HELIX RES INST.

Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI, 2001-318749/34.
XX

PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -

Claim 8; SEQ ID 14558; 2537bp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC polynucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets are useful for antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 2272 BP; 609 A; 509 C; 585 G; 569 T; 0 other;

XX Query Match 9.5%; Score 329; DB 22; Length 2272;

XX Best Local Similarity 51.8%; Pred. No. 6,1e-72;

XX Matches 871; Conservative 0; Mismatches 791; Indels 18; Gaps 5;

```

Qy 1086 TTGTTTACCAATATATACATTAAGATCAGAGTGTATGATGATGATGATGATGATG 1145
Db 34 TTGTTATCATCTGATATTAACATGATGATGATGATGATGATGATGATGATGATGATG 93
Qy 1146 ATCCCAAGAGACCTTTAAGAAAGCCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1205
Db 94 CTCCAAGAGATGCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 153
Qy 1206 GGAAGCATATCAACCAAGATACCGACTTGAAGAGCGCTGCTGCTGCTGCTGCTGCTG 1265
Db 154 GCAAAATTTATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 213
Qy 1266 ---AGAGAGCGCGGCGCTGCGGCGGAGCACTGCGAGGCGCTGCTGCTGCTGCTGAGC 1322
Db 214 GTGAGAGACAGATATATCATGAGGATGCTGCTAAGAGGGAATCTGCTGCTGCTGAGC 273
Qy 1323 TCTGCTATCTTAACAGCTTAAATGCTGATATGATGATGATGATGATGATGATGATGAT 1382
Db 274 TTTCTTTATGACCGGATGCCAGAGATGAAAGAGCTTCAAGAGCATTAAGATTT 333
Qy 1383 TAGCGCTTATACAGAGCTAATCTCAGAGCAAGGAGCGCTGATGATGATGATGATGATG 1442
Db 334 TGGCTCAGCAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 393
Qy 1443 ATTAATTCATTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1502
Db 394 AAGAAATTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 453
Qy 1503 ATTCGAATTAATGCTGCTGCTGAGAGAGATTTGCAAGAGAGATTCACCAAGAGTG 1562
Db 454 AAGAGATGATCATTAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 510
Qy 1563 GAAAGATTTGATTAATCAATCAATTTGCAATGATGATGATGATGATGATGATGATGAT 1622
Db 511 AATATATCTGTTATATCAATTCAGAGACTAATCTGTTAAGAGATGATGATGATGATG 570
Qy 1623 CATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1682

```

Db 571 CTTCATCTTGACTATCCCATGATCTTCTGGCATTCTTTTACCCAAAGAGCAATGG 630
QY 1683 AAGCAGCAATTCATTGATACAAAATCTATTTAAAGTTACACAGCCATGGGCATGCAAA 1742
Db 631 ACCAGGCTCGAAGACTGGTCAACATGTTGGAGAAGATAGCGGCCCAATTGGCATGCGTA 690
QY 1743 TGAGAAAGCAATAATGATTTGAAGT---GGATGACAGAACTGAAGCCTACTTAAAGATCT 1799
Db 691 TGAGCCCAACCGGCTGGGTTGAATTAAGGATGACCGAATAGAGACTTATGTGCAAGACA 750
QY 1800 TACAGCAAAAGGTACAGCAG---ACACCCAGATAGTGTCTGCTGTTGTCAAGTA 1853
Db 751 TCAATCCAGTTAGAGCTGAGGGGAAGATACAGATGTTGTTTGCATCATCATGGGCC 810
QY 1854 ATCGGAAGGACAAATACGATGCTATTAAAAAATACCTGTGTACAGATTGCCCTACCCCAA 1913
Db 811 CACGTGATGATCTCTATGGGGCCATCAAGAAGCTGTGTGTGAGTCCCCAGTGCCT 870
QY 1914 GTCAAGTGTGGTGGCCGCACTTATAGGCAACAGCAAACTGTCTAGGCCATTTGCTACAA 1973
Db 871 CCAGGTTGTCAATGTTGCAACCAATTTGGTCAGCCCAACAGGCTTCGGAGTGTGGCCCA 930
QY 1974 AGATTGCCCTACAGATGAATCAAGTCAAGATGGAGGAGCTCTGGAGGGTGGACATCCGCC 2033
Db 931 AGATTTTACTTCAGATTAATCTGAAATTTGGTGTGAGCTCTGGGAGTGGATATTCCTC 990
QY 2034 TGAAGCTCGTATGATCGTTGGCATCGATTGTTTACATGACATGACAGCTGGCGGAGGT 2093
Db 991 TGAACAGTTAATGGTGTATCGGGATGATGTTTACCATGACCCAGTACAGGATGCGCT 1050
QY 2094 CAATCGAGGATTTGTTGCCAGATCAATGAAGGATGACCCGCTGGTCTCACGCTGCA 2153
Db 1051 CCGTGTGTGGCTTGTGGCAAGCATCAATCTCACCTCAAAAATGGTATTCCCGGGTGG 1110
QY 2154 TATTTTCAGGATAGAGCAGGAGCTGTAGATGGCTCAAGCTCTGCCCTGCAAGCGGCTC 2213
Db 1111 TGTTTCAGATGCGCATGAGGATGTTGGACAGCTGAAAGCTATGCTCTGCTGGGCTCT 1170
QY 2214 TGAGGCTTTGGAATAGCTGCAATGATGATACATCCAGCGGATCATCGTGTACCGCGATG 2273
Db 1171 TAAAAAGTTTATGAGGTGAACCACTGTCTACAGAGAAGATTTGGTGTACCGTATG 1230
QY 2274 GGTAGGAGCGGCTGAGCTGAAACACTGCTGTAACAGAGTGCACAGTTTGTGAT 2333
Db 1231 GAGTGTCTGATGGCCCACTGAAGACAGTTGCCAAGTATGAGATTCTCAACTACAGAAAT 1290
QY 2334 GTCTAAATCCATTGTTAGAGGTTTCAACCCCTAGANTACGCTAATTTGGTGAAGAAA 2393
Db 1291 GTTTGAAGCTTTTGAGA---ATTATCAGCCCAAGATGGTGTGTTGTAGTTCAGAAGA 1347
QY 2394 GAGTGAACACCAAGATTTTGTCTCAGTCTGGAGGAGACTTCAGAAATCCACTTCCCTGAA 2453
Db 1348 AAATCAGTACTAATCTATATCTGGTGTCTCTCAGAACTTTGTAACCTCCACTCTGGAA 1407
QY 2454 CAGTATTCAGTATGAGGTTACAGACCAAGATGGTATGACTTTTATTCGTGAGCCAGG 2513
Db 1408 CTGTGTAGATCATACAATAAACAAGCTGTGAGTGGGTGATTTCTATCTCTTGTGCCATC 1467
QY 2514 CTGTGAGAAGTGTAGTGTCTTCCACACATTTACAATGTCTATGACAAACAGCGGCC 2573
Db 1468 ATGTACCGCAGGGCTGTGCATCTCTACGCCATTTATGTCTGTGTTCTCAACCCGCAACC 1527
QY 2574 TGAAGCCAGACACAPACAGCGCTTGACCTACAAGCTGTGCCACATCTATTACAACTGGC 2633
Db 1528 TGAGCCCTCATATGACAGAGGCTGACTTTCAAACTGTGCCACATGTAAGTGAATTTGGC 1587
QY 2634 CAGGTGTCAATCGTGTCTCTGCTGCTGAGTACGCCCAAGCTGGCTTTCTTGTG 2693
Db 1588 CTGGCAACCATCAGAGTTCCAGTCTCTTGAAGTATGCCCAAGCTAGCTTTCTCTGTGAG 1647
QY 2694 GCCAGATTTACAGAGAGCAAAATCTGTCACTGTCAACCGCCCTTTTACTACCTCTAAC 2753

Db 1648 GACACATCTTGCATCATGGACCAGCCATCCAGCTGTGCGAGAACCCTGTTCTCTCTGTGAC 1707

RESULT 9

ABL10571

ID ABL10571 standard; cDNA; 2838 BP.

XX ABL10571;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 26195.

XX Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX W0200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX P-PSDB; ABB66468.

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

XX Claim 1; SEQ ID NO 26195; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 2838 BP; 783 A; 659 C; 682 G; 714 T; 0 other;

Query Match 8.9%; Score 310.2; DB 23; Length 2838;
Best Local Similarity 49.1%; Pred. No. 3.5e-67;
Matches 1133; Conservative 0; Mismatches 1130; Indels 46; Gaps 10;

QY 476 AAAACAGTCTTTCAGGCATTATAGTAAAGTTAAGCACTAACCACTTCCGGCTGACATCC 535

Db 431 AAGAAAGAGTGGTGGCACTCATATTACCGTCGAGGCAAACTATTTAAGGTATTAAAG 490

QY 536 CGTCCCGCAGTGGCGCTTATATCAGTATCACTTGAATTAACCCCACTGATGGAAGCCAGA 595

Db 491 CGTCCAAACTGGACCATCTACCACTACCGCTCGATTTTACGCTCATGTGGAGGCTACA 550

QY 596 AGACTCGTTCAGCTCTTCTTTTCAACGAGATCTAATTTGGAAGTGTATGCTTTT 555

Db 551 CGACTGCGAGGCTTTCTTGTATGAACATAAAGGATCTCTGGGC--GGCTACATCTTTG 608

QY 656 GATGGAACGATATATTTTACCTAAAGACTACAGAAA-----AGTTTA 701

Db 609 ACGAACAATATGTTTTCATCAATCACTTCAAGCTTCAAGTAGCCCCCTATGTTT 568

QY 702 CTGAAGTTTGTAGTAAAGCCGGAATGAGAGATGTGAGATTAAGATCACTTTAAACA 761
 DB 669 TGGAACTGGTTACGAAGAGTCTGCTGGGAAAAACATTTGAATTAAGATCAAGGCTGTG 728
 QY 762 ATGAACCTTCCACTACATCAACCACTTGTGCACTTATATATATATTTTTCAGAGAGC 821
 DB 729 GATCTGTGCAATCTACGAGATGCGGAGAGTTTCAAGTCTTATCTACTCTGCGAGGG 788
 QY 822 TTTTGAATATCAATGATTTGCAACAAATTTGAGCAAAATTTATATTAACCAATGACCCA 881
 DB 789 CCATGAGAGGCTTAAGCTTGAAGTGTCTGCGCTACTACTACATCTCAAGCTAAGA 848
 QY 882 TTGATATTTCAAGTCAAGTGTGATTTGGCTGCTGCTTCACTTCACTTCCATCTTCACT 941
 DB 849 TTAATTTGAGAAATTTCCGCAATGCAATTAATGCTGCTGCTATCAAGCTTCAATCCGAGC 908
 QY 942 ATGAAAAAGCATCATGCTCTGCACTGAGCTTAAAGCTCAATAGTCTTGAAGTGAAGT 1001
 DB 909 AGGAAATGATATATTTACTTGTCTCGAGATATGCAACAGGTTATGCGAACTGAGACT 968
 QY 1002 TTTTGAATTTCAATGTTCACTTTTATCAATCAAGAAACATTAATTTCAAGAAACAG 1061
 DB 969 TGTACAAATATTTTGTCCAGTCTATTCGAGAGTGAAGATTAAGATCAATTTAAGC 1028
 QY 1062 TTTCCAAAGAAATTAATAGTGTAGTTGTTCTTACCAAGTATTAATTAAGATCAATCAAG 1121
 DB 1029 GT-----GCACTTAATGGGTATGTTATTAATTAATTAATTAATTAATTAATTAATTA 1082
 QY 1122 TGGATGATATTTAGCTGGAGCAAGAAATCCCAAGAGCACTTTAAGAAAGCCAGGCTGTG 1181
 DB 1083 TTGACGATGTCGACTTTCATGAGCGCAATGTCGCAAAATTT---AAACCAATGAGCGGTG 1139
 QY 1182 GGGTCAAGCTTTTAAAGATCTACAGAAAGCAATTAACAAGATTAACCAAGTCAAGCTTGAAGC 1241
 DB 1140 AATTTTGTGATGATGATTAATTAAGAAAGCAATTAACAATTAATTAATTAATTAATTAATTA 1199
 QY 1242 AGCCTGTCTTGTGTCAGCCAGCCCAAGAGAGGCGGAGCTTGGGAGGACATCTGCAAGGCGC 1301
 DB 1200 AACCTGTGTCATGTCCTC-GTCCGACAGCAAAATTTTGTGATGAGAAATGACAGGCGC 1258
 QY 1302 CTGCACTGCTCATCTCTGAGCTCTGCTATCTTAAGGCTTAACTGATTAATTAATGCTTAATG 1361
 DB 1259 ATTA--ATGATCAATCTCCGAGCTGAGCAAGGCTTAAGGAAATGACGAGCTTATGCGCTG 1316
 QY 1362 AATTTAAGCTGATGAAGACTAGCGCTTCAACAAGATTAATCTCAAGAGCAAGGAGCAGC 1421
 DB 1317 ACTTTAGAGCTTGAAGGCGCATGAGTGAACATACAGGCTGATCAAGATGCTGCAATCG 1376
 QY 1422 GTGAAGTGGAGAGCTATGATTAATCAATTAACAAAGATTAATGTTCAAGAGGAGCTTC 1481
 DB 1377 AAGGCTTGCATGTTTCAACAGCGCTTAAAGTCTGTTAAACAGAGTGAAGAGCCCTTA 1436
 QY 1482 GAGACTGGGCTTGAAGCTTGAATCCAACTTACTCTCTCAAGAAAGATTTTGCAGAA 1541
 DB 1437 AGTCTTGAAGCATGAGCTGAGCTGCGCTTGTGAGAGATTCACAGCGCGTGTGCAC 1496
 QY 1542 CAGAAAGATTTACCAAGGTGAGAAACATTTGATTAACATCAATTTGAGATTTGAGT 1601
 DB 1497 CGGAAAAATATTTATTTGCGCAACAAAAATATTCGATGCGAGCTGCGCGCAATTTGA 1556
 QY 1602 CCAAGAAACCAAGAGGTGACATTAATTAAGTTAAGCACTAAGATTAATGAGCTTGA 1661
 DB 1557 CCAATGATTTGCAAGCTGTTGATGTTTAAAGGTGATTAATTAATGAGTGAATGAG 1616
 QY 1662 TCTATTAAGGAAATTAATTAAGAGCAAGCAATTAATTAATTAATTAATTAATTAATTA 1721
 DB 1617 TCACACCGAGTGAATATGCGAGAGACCCAGAAATTTGTGCAAGTGTGACATCCGACAG 1676
 QY 1722 CACGAGCATGGGCAATGATGAAGAAAGCAATATGATTAAGT---GATGACAGAA 1778
 DB 1677 CCAGCAGCATGAAGTGAATCTGCAATCAATTAATTAAGAAATTTCCGATGACCGCA 1736

QY 1779 CTGAAGCTTACTTAAGAGTCTTAACAGCAAAAGTCAACAGAGACACCAAGATAGTTGCT 1838
 DB 1737 AGGCACTTAATCTCCAAAGGATGACACAGCGGCAAGCTTAACATGATCAAGATGATGAG 1796
 QY 1839 GTCTGTGTCAAGTAAATCGGAAGAGCAATTAAGATCTTAAATTAATTAATTAATTAATTA 1898
 DB 1797 TTGTCAATGAGATCTCCAAATGAAGAAATTAATGATGATTAATTAATTAATTAATTAATTA 1856
 QY 1899 ATTGCCCTACCCCAAGTCAAGTGTGT-----GTGGCCGAACTTTAGGCAAAACAG 1948
 DB 1857 ACAGACCGGTGCGCATGCGAGGTGTGACACTAAAGTTATTCGCGCTTCAGACAGCAAAAC 1916
 QY 1949 CAAACTG--TCATGAGCTATTTCTTAACAAAGATTCCTTACAGATGAATCTCAAGATGAGG 2006
 DB 1917 CAATGGGCTGATGTGATGATGCGCAAAAGTGGTTATTCAAATTAAGAGCCCAATTAATGAG 1976
 QY 2007 GAGAGCTCTGAGAGGTGAGATTCCTCCCTGAAGTCTGATGATGATGATGATGATGATGATG 2066
 DB 1977 GAGCTCCCTGGCAGGTAGATGATCCCTCCAGGCTGATGATGATGATGATGATGATGATGAT 2036
 QY 2067 ACATGACATGACAGGCTGGGCGGAGGTCAATGCGAGGATTTGTGTCAGATCAATGAAG 2126
 DB 2037 GCAATTCACCGAAGAAATTAAGAAACATGATTAATGAGGCTTTGTTGCAACATGAGACCA 2096
 QY 2127 GATATGAC---CCGCTGATCTCAAGCTGCAATTTTCAAGATGAAGAGAGAGCTGTAG 2183
 DB 2097 AGGAGTCTTTCCGCTACTTCTCCACCGTAAAGCAACATTAAGGCGCAAGAGTGTCCG 2156
 QY 2184 ATGGGCTCAAAAGTCTGCTGCAAGCGGCTCTGAGGCTTGAATGATGATGATGATGATGATG 2243
 DB 2157 AGCAGATGTCGAGTGAACATGAGGCTGCGCTGAGGTGATGATGAGAGCAACCGTTCTT 2216
 QY 2244 TGCCAGCGGATCATGATGATGATCCGAGATGAGCTGATGAAGAGCGGCACTGAAACACTG 2303
 DB 2217 TGCCAGAGGCAATCTCTTCTTCTGCGAGCGGTGTTGCGATGATGATGATGATGATGATG 2276
 QY 2304 TGAATCAAGAGTGAACAGATTTTGAATTTGATTTCAATTTGATGATGATGATGATGATG 2363
 DB 2277 TAAACAGCGAGTGAACACCGTAAAGAGAGGCTGACGAAATTTTCAATCAAGCTGAGCA 2336
 QY 2364 CTGAGATTAAGGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2423
 DB 2337 AACAGAGGCTGTGCAATGATTAATTAATTTGATTAATTTGATTAATTTGATTAATTTG 2396
 QY 2424 GAGGAAGACTTCAAGATTCACCTCTGGAACAGTTATGATGATGATGATGATGATGATGATG 2483
 DB 2397 TTAATGAGCATGCAACCGATTTCCGAGCACTGATGATGATGATGATGATGATGATGATG 2456
 QY 2484 AATGATATGACTTTTATCTGAGGCAAGCTGTGAGAAATGATGATGATGATGATGATGATG 2543
 DB 2457 AGCGTACGACTTCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2516
 QY 2544 ATTAACATGATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2603
 DB 2517 GCTACATATGATTTTCTGACACATGAGGACTTAAAGCGGATTAAGCTGAGATGCTCTCT 2576
 QY 2604 ACAAGCTGTGCAATCTATTAACAATGAGCTGATGATGATGATGATGATGATGATGATGATG 2663
 DB 2577 ATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2636
 QY 2664 AGTACGCCCAAGCTGCTTTTCTTGTGGCCAGATATTCACAGAGGCAATCTGT 2723
 DB 2637 ATTAAGGCCCAATATGAGCTTCTCTGATGCGAATTCATTAATCTGAGGCTTCAAGCAG 2696
 QY 2724 CACTGTCAAACCGCTTTACTACTCTAA 2752
 DB 2697 GACTGCAAGATCAATGATCTTTTGTAA 2725

RESULT 10
 AA213404
 ID AA213404 standard; cDNA; 300 BP.
 XX

AC AAZ13404;
 XX 12-OCT-1999 (first entry)
 XX Human gene expression product cDNA sequence SEQ ID NO:873.
 XX Human; gene; gene expression product; diagnosis; therapy; probe;
 KW detection; mapping; tissue typing; profiling; forensic; cancer;
 KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
 XX Homo sapiens.
 XX WO9938972-A2.
 XX 05-AUG-1999.
 XX 28-JAN-1999; 99WO-US01619.
 XX 03-APR-1998; 98US-0080666.
 XX 28-JAN-1998; 98US-0072910.
 XX 24-FEB-1998; 98US-0075954.
 XX 31-MAR-1998; 98US-0080114.
 XX 03-APR-1998; 98US-0080515.
 XX (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
 PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
 PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
 PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
 PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
 XX WPI; 1999-494092/41.
 DR Novel human genes and their expression products which are
 PT differentially expressed in different cell types
 XX Claim 1; Page 862-863; 2479pp; English.
 XX The present invention describes a library of human polynucleotides
 CC comprising the sequences given in AAZ12532 to AAZ17779. Also described is
 CC a method of detecting differentially expressed genes correlated with the
 CC cancerous state of a mammalian cell, comprising detecting at least one
 CC differentially expressed gene product in a test sample from a cell
 CC suspected of being cancerous, where the gene product is encoded by one
 CC of the 5248 polynucleotide sequences given in AAZ12532 to AAZ17779. The
 CC polynucleotides can be used as a source of primers and probes, which can
 CC be used for a variety of purpose, e.g. detection of expression levels,
 CC mapping, tissue typing or profiling, forensics, genetic analysis and
 CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
 CC can be used for raising antibodies for experimental, diagnostic and
 CC therapeutic purposes. The polynucleotides may also be used to construct
 CC arrays for diagnostics (which may be used to determine function of an
 CC encoded protein); and to detect differences in expression levels between
 CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
 CC identify a genetic predisposition or susceptibility to a disease such as
 CC cancer). The polynucleotides of the invention are especially used in the
 CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
 CC and lung cancer. The polynucleotides can also be used to screen for
 CC peptide analogues and antagonists.
 XX Sequence 300 BP; 75 A; 70 C; 91 G; 64 T; 0 other;
 SQ Query Match 8.2%; Score 286.4; DB 20; Length 300;
 Best Local Similarity 99.3%; Pred. No. 1.1e-61;
 Matches 298; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1909 CCCAAGTCAAGTGTGTGGTGGCCGCAACCTTAGGCAACAGCAAACTGTGATGGCCATTCG 1968
 Db 1 CCCAAGTCAAGTGTGTGGTGGCCGCAACCTTAGGCAACAGCAAACTGTGATGGCCATTCG 60
 QY 1969 TACAAGATTGCCCTACAGATGCACTGCAAGATGGAGGAGAGCTCTGGAGGTGGACAT 2028

Db 61 TACAAGATTGCCCTACAGATGCACTGCAAGATGGAGGAGAGCTCTGGAGGTGGACAT 120
 QY 2029 CCCCTGAAGTCTGTGATGATCGTTGGCATCGATTTCATCATGACATGACAGTGGCG 2088
 Db 121 CCCCTGAAGTCTGTGATGATCGTTGGCATCGATTTCATCATGACATGACAGTGGCG 180
 QY 2089 GAGTCAATCGCAGGATTTGTTCCAGCATCATGAAGGATGACCCGTTGTTTCACG 2148
 Db 181 GAGTCAATCGCAGGATTTGTTCCAGCATCATGAAGGATGACCCGTTGTTTCACG 240
 QY 2149 CTGCATATTTCCAGATAGAGGACAGGAGCTGGTAGATGGCTCAAAGTCTGCCTGCAAGC 2208
 Db 241 CTGCATATTTCCAGATAGAGGACAGGAGCTGGTAGATGGCTCAAAGTCTGCCTGCAAGC 239
 RESULT 11
 ABL10535
 ID ABL10535 standard; cDNA; 3524 BP.
 XX ABL10535;
 AC ABL10535;
 XX 26-MAR-2002 (first entry)
 DT 26-MAR-2002 (first entry)
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 26087.
 DE Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX Drosophila melanogaster.
 XX WO200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US09231.
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR P-PSDB; ABB66432.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PS Claim 1; SEQ ID NO 26087; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 3524 BP; 1053 A; 741 C; 761 G; 969 T; 0 other;
 SQ Query Match 7.5%; Score 259.8; DB 23; Length 3524;
 Best Local Similarity 47.3%; Pred. No. 1.7e-54;
 Matches 1085; Conservative 0; Mismatches 1179; Indels 29; Gaps 9;
 QY 470 GAATCAAAAACAGGTTCTTCAGGCATTATAGTAAGGTTAAGCACTAACCACTTCCGGCTG 529

Db 342 GATCCAAAGAGGAAACCGATGCGGCCGATCATGCTGACAGAACGAACTTTTCCGATTA 401
 Qy 530 AATATCCCTCCCAAGTGGCCCTTATATCATGATTCATTTGATTAACCCACTGATGGA 589
 Db 402 AAAACCCAAACCCGAAATGGGAGTCGTTATTAACAGTGGAGTTGAGCCAGACATGAG 461
 Qy 590 GCGAGAAAGCTCCGTTAGCTCTTCTTTTCAACAGAAAGATCTAATTTGAAAGTGTAT 649
 Db 462 AATCTCGTGTCCGATGAGGATTTTTCGATTCATGCTAACCTTCTGGGATCAGGCTAT 521
 Qy 650 GCTTTTGTAGGAAGCATATTTATTTTAAAGCTAACAGCAAAAGGTTATCTGAAGTT 709
 Db 522 CTATTCGATGAGCTGCAATGTTTCAACACAGAAATTCAGCAAGAAATCAAGTGTCTC 581
 Qy 710 TTTTAGTAACCCCGAATGGAAGATGTAAGATTAAGATCACTTTAACAATGACTT 769
 Db 582 AGCGAAAGTCGAAAGTCGACATTTGAATACAAAGATTCATTAAGTGTGTTGATTCATA 641
 Qy 770 CCACCTACATCAACCACTGTTTTCAGTTCTAATAATATTTTTCAGAGGCTTTTGAAA 829
 Db 642 TGTGTGCTGAGCCCGCTTTTTCGAAATCTTAATCTAATATGTCGCGCTCCATGAA 701
 Qy 830 ATCATGAATTTTGCACAAATTTGAGCAAAATTTATTTATTAACCCCAATGACCCAA 889
 Db 702 GGCCTAAATTTTGAATTTAGTTGGCCGTAATCTTTTGAATCCCGAGCTAAGATTCGAA 761
 Qy 890 CCAAGTCACAGTTGTTGATTTGGCCGCTTCACTACTTCATCCCTCAGATGAAAC 949
 Db 762 AGGAGTTCAAAATGAGCTATGCGGCTATGAGACATCAATTTGTGAGACGAAABA 821
 Qy 950 AGCATATGCTGCACTGACGTTAGCCATTAAGTCTTGAAGTGAAGCTGTTTGTGAT 1009
 Db 822 GAAATTTTATTTGGGACCGAAATTAATCAAAAGTATGAGGACGAGACGATCAACAC 881
 Qy 1010 TTCAATGTTCACTTTTATATCATCAAGAAAGCAATTAATTTCAAGAAACAATTTCCAA 1069
 Db 882 ATATGCGAGCTGTCTCAACATCCGCTGCTCATCAGAGCAAGATGAGGCT-----A 935
 Qy 1070 GAAATTAAGTGTAGTTGTTCTTACCAAGTATTAATTAAGCATACAGAGTGTAT 1129
 Db 936 AATGTTTGGACTGATTTGTTCTTACGATTAACATTAAGAACTTATGCTATCATGAT 995
 Qy 1130 ATTGACTGGACAGAAATCCCAAGACACCTTTAAGAAAGCCGAGCTCTGCGGCTACG 1189
 Db 996 GTGACCTTTGGAACAACTCCGAATCAACAT---CAGTTGCAAGGTTAGAGTATCAGT 1052
 Qy 1190 TTCTTGAATTAATCAAGAAAGCAATTAACCAAGAGATCACCGACTTGAAGACGCTGTC 1249
 Db 1053 TTGCTGAAATTAATCTCACTAATATTAATATGCAATTCGCAACCAATCAAGCTGTC 1112
 Qy 1250 TTGCTGACGAGCCCAAGAGAGCGGCGCTGCGGCGGACACTGCAGGCGCTGCGCATG 1309
 Db 1113 CTGATCTCCAAA---AATAGGACAGAGCTCTTAATACTTAACGCTAGCAATTAAGTGA 1169
 Qy 1310 CTCAATCTGAGCTGCTATCTTACAGTCTTAATCTGAATTAATGCGTATGATTTTAA 1369
 Db 1170 CTAAATCTGAGCTGCTGCGAGTGAAGTGGCTCAATGCGGAGTGGCTCAAACTTTGAG 1229
 Qy 1370 GTGATGAAGAACTTAAGCGCTTACAGAACTAATCTCAAGAGAAAGCGAGCTGAAGT 1429
 Db 1230 CTATATGCTGCGATGAGCAAGTATTAAGCAATGAACCCCAACAGCACTGATCAATG 1289
 Qy 1430 GGAAGCTCAATTTGATTCATTTCAATTAATAAGATTAATGTTCAAGAGGCTTCAGACTG 1489
 Db 1290 CGCGCTTTTAAACACCGTTTAAACAAACATCTCAGAAAGTGTGAAGGCTTGAAGACTG 1349
 Qy 1490 GGTTTGAGCTTTGATTCAACTTACTGCTCTTCAAGAAAGATTTTGCAGAAAGAAAG 1549
 Db 1350 AACATGAAGTGAAGAAAGTCAAGAAAGTCAAGAAAGTCAAGGCGGATTAATGAGACG 1409
 Qy 1550 AT-----TCACCAAGTGAAGAAACATTTGATTAATCAATCAATTTGAGATTTGTC 1605
 Db 1410 ATCGTGTTCATTAATGAAGGTTCTGCTGAGAAACGCTGATTTGCAAAAGGCACTTC 1469

Qy 1606 AGAAAGAGAGTGCACATTAATTAAGTTAAGCACTAGATTAACGCTGTTGATCTTA 1665
 Db 1470 AGAGACCAAAAGATGCTTACCACTCCAGAGATGCG-CTCGATCGTTGGCTGTCACTCC 1528
 Qy 1666 TACCGGAAGAAATTAAGAGACGCAATTCATTAATTAAGAAATCTAATTTAAAGTTAC 1725
 Db 1529 GCGCAAGAGAAATTTCCATGAATCCGAACTCACTTGAATCTTGTATTAAGAGAGCTG 1588
 Qy 1726 AGCATGCGCATCAATGAGAAAGC---AATATGATTTGAAGTGAATGAGCAAGATCTGA 1782
 Db 1589 TGAATGAGGCTTGAATTTGAAGAGCCCGAGAAATTCATTAATTAAGATGATGAGCTG 1648
 Qy 1783 AGCTTACTTAAGCTTTACAGAAAGTCAAGAGACACCCAGATATGTTGTCTGTCT 1842
 Db 1649 AACTTATGAGAGCAATGATGATTTGTGCGCTCAGATCCCAACTAATATTAATGAGCT 1708
 Qy 1843 GTTGTCAAGTAATGAGAGCAATTAAGATGCTATTAATAAATACCTGTATCAGATTTG 1902
 Db 1709 GGTACCAATGATTAACCCCAAGATATCTCATATCAAAAGAGAGATACGTTACAG 1768
 Qy 1903 CCTAACCCCAAGTCAAGTGTGTTGCGCCGAACTTAAAGCAACAGCAACCTGATGCG 1962
 Db 1769 GCGGTGCACTCA---AGTTGTGACCTTTAAAGACCAAGAACCGTATGAG 1825
 Qy 1963 CATTTCAAAAGATTTGCTCAAGATGAATCTGAAGATGGAAGAGCTCTGAGGCT 2022
 Db 1826 CATTCACCAAAATAGCAATCAATCAATGAAATTCAGATTTGAGATTAACACCTGATGAT 1885
 Qy 2023 GGAATATCCCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2082
 Db 1886 GGAATATCCCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1945
 Qy 2083 TGGCGAGAGTCAATCCAGAGATTTGTTGCAAGATCAATGAAGAGATGAACCGCTGTT 2142
 Db 1946 TCGGAAGAGGCTTACGAGAGATTTGATTTGCTCAATGATCTACAGCAAAATTCACAGTA 2005
 Qy 2143 CTACGCTGATATTTCAAGATTAAGAGACAGAGCTGTGATGAGTGTGAGTGTGAGTGTGAG 2202
 Db 2006 CTTCAGACACATCAAGAGTGTGAGAGCTGTTGATGTTGCTGCTCAACCTTTGCGCAT 2065
 Qy 2203 GCAAGC---GGCTGAGGCTTGAATGAGTGAATGATGATGATGATGATGATGATGATGAT 2259
 Db 2066 GATAGCAAGGCTTGGGCTCAATTAATCAATGATGATGATGATGATGATGATGATGATGAT 2125
 Qy 2260 GGTGTACGCGATGCGTGAAGAGAGCGAGCTGAAGAAACCTGTGATGAATGAAGTGC 2319
 Db 2126 ATTTTATGAGAGCGTGTGAGCTCGGCTCTTAAGCAGCTTTTGAATTTGAAGTCA 2185
 Qy 2320 ACGTTTGTGATTTGTTAATCAATTTGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2379
 Db 2186 GGAATCAATGAGAAATGGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 2245
 Qy 2380 TGTGTGAAGAAAGAGTGAACCAAGATTTTGTCTGAGCTGAGAGAAAGCTTCAAG 2439
 Db 2246 AGCT---TATATTTGTGATTAACAGATCAATGAACAGCGCTTCTTCTCAAGCAAAA 2302
 Qy 2440 TCACTTCTGGAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2499
 Db 2303 TCTTCCGCTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2362
 Qy 2500 TATGTGAGCGAGCTGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2559
 Db 2363 TCTGTGTCTGCAACAAAGTTTGTCAAGGATCAAGTGTGCGAGCAGCTCAATGATTTT 2422
 Qy 2560 TGAACAAGCGGCTGAAGAGCAACATACAGAGCTTGAATCAAGAGCTGTGCAAT 2619
 Db 2423 TAGCAGCATGAGTCTCTCAACGAGAAATTAAGCAAAATTAAGTGAAGTGTGCACTT 2482
 Qy 2620 CTATTAAACATGAGCGAGTGTCAATGTTGTTCTGCTCTTGTGAGTGAAGCGCAAGCT 2679
 Db 2483 GTACTAACATTTGTTGCGGACACACAGAGTGCAGCAAGTGTGCAAGTGAAGAGCT 2542

Qy	2680	GGCTTTTCTTGTGGCAGAGTATTCACAGAGAGCAATCTGTCACTGTCAACACGCCT	2733		
Db	2543	AGCTACCTCGTGGTACGAACCTTGCACTCTATTCCGAAAAACGCTCGAAAGAAGTT	2602		
Qy	2740	TTACTACCTCTAA	2752		
Db	2603	TTATTATCTATAA	2615		
RESULT 12					
ID	AAA07586	AAA07586 standard; DNA; 3047 BP.			
XX	XX				
AC	AAA07586;				
XX	XX				
DT	29-AUG-2000	(first entry)			
XX	XX				
DE	Drosophila piwi gene.				
XX	XX				
KW	Piwi family protein; piwi; miwi; hiwi; gene therapy; tissue dystrophy;				
KW	anaemia; immunodeficiency; male infertility; Drosophila; ds.				
XX	XX				
OS	Drosophila sp.				
XX	XX				
FH	Key	Location/Qualifiers			
CDS	84..2615				
FT	/*tag= a				
FT	/product= piwi				
FT	/transl_except= (pos:120..122; aa:Xaa)				
FT	/transl_except= (pos:399..401; aa:Xaa)				
FT	/transl_except= (pos:2436..2438; aa:Xaa)				
FT	/note= "Xaa= Leu Or Ile"				
XX	XX				
PN	WQ200032039-A1.				
XX	XX				
PD	08-JUN-2000.				
XX	XX				
PF	03-DEC-1999;	99WO-US28764.			
PR	04-DEC-1998;	98US-0110901.			
XX	XX				
PA	(UYDU-) UNIV DUKE.				
XX	XX				
XX	Lin H;				
XX	XX				
DR	WPI; 2000-412085/35.				
DR	P-PSDB; AA90233.				
XX	XX				
PT	Piwi family nucleic acids, polypeptides, and antibodies, useful in gene				
PT	therapy of diseases such as cancer and in various research and				
XX	diagnostic applications -				
PS	Claim 19; Page 171-175; 201pp; English.				
XX	XX				
CC	This sequence encodes the Drosophila piwi family protein, designated				
CC	piwi. The piwi family nucleic acids and polypeptides are used in gene				
CC	therapy of diseases such as cancer and also in various research and				
CC	diagnostic applications. The sequences can also be used to treat				
CC	tissue dystrophy, anaemia, immunodeficiency, and male infertility.				
XX	XX				
SQ	Sequence 3047 BP; 936 A; 664 C; 672 G; 772 T; 3 other;				
Query Match					
Best Local Similarity 7.3%; Score 253; DB 21; Length 3047;					
Matches 1080; Conservative 0; Mismatches 1184; Indels 29; Gaps					
Qy	470	GAATCAAAACAGGTTCTTCAGGCATTATAGTAAGTTAAGCACTAACCATTTCCGGCTG	529		
Db	342	GTATCAAGAGGAAACGATGGGTCGCCGTATGTCGAGACGAACTTTTTCGANTA	401		
Qy	530	ACATCCCGTCCCGATGGGCCCTTATATCAGTATCACAATTGACTATAACCCACTGATGAA	589		
Db	402	AAAAACAAGCCGGAATGGCGATCGTTCAATTATCACTGGAGTTTGTGGCCACCATCGAG	461		

QY 1666 TACGCGAAGAAATTAAGAGAGCAATTCATTGATACAAATCTATTAAAGTTACACC 1725
 DB 1529 GCGCGAAGAAATTCCTCAAGATCTCGAATCTTGTATAGAGAGCTAG 1588
 QY 1726 AGCATGGGATGCAAAATGAGAAAGC--ATAATGATGAAGTGAAGACAGACTGA 1782
 DB 1589 TGGAAATGGGTCTTAAGAAATTCGAAGCCCCAGGAATTCATTAATTAATGATCCACTGG 1648
 QY 1783 AGCTTACTTAAGAGCTTACAGCAAAAGTTCACAGACACCCGAGATGTTCTGCT 1842
 DB 1649 AACTTATGAGAGCAATGATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1708
 QY 1843 GTTGTGAAGTATCGAAGAGCAAAATACGATCTATTAAAAATACCTGTGTAAGATTG 1902
 DB 1709 CGTACCAATGATTAAGCGGAAAGATATCTCAATCAATCAAAAGAGAGATCGTTGACG 1768
 QY 1903 CCGTACCCAGTCAAGT 1962
 DB 1769 GCGGATGCAACTCA--AGTTGTGACCTTTAAACGACCAAGAACCGTATGAG 1825
 QY 1963 CATTGCTACAAAGATTGCTTACAGATGAATGCAAGATGAGAGAGCTTGAAGGT 2022
 DB 1926 CATTGCTACAAAGATTGCTTACAGATGAATGCAAGATGAGAGAGCTTGAAGGT 1885
 QY 2023 GGAATCCCTGAAAGCTGTGATGATGCTTGGCATGATGCTTGAATGATGATGATGAT 2082
 DB 1886 CGAATACCTTGTCCGAGCTGATGACAAATGCTTGAATGATGATGATGATGATGATGAT 1945
 QY 2083 TGGGCGAGAGCTCAATCCGAGATTTGTCAGATCAATGAAGAGATGACCGCTGAT 2142
 DB 1946 TCGAAGAGGCTCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2005
 QY 2143 CTCAAGCTGATATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2202
 DB 2006 CTTCAGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2065
 QY 2203 GCAAGC--GGCTGTGAGAGCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 2259
 DB 2066 GATACGAAAGGCTGCGCAATATCAATGATGATGATGATGATGATGATGATGATGATGAT 2125
 QY 2260 CGTGTACCGGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2319
 DB 2126 ATTTTATGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2185
 QY 2320 AAGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2379
 DB 2186 GGAATATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2245
 QY 2380 TGTGTGAAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2439
 DB 2246 AGCT--TATATGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2302
 QY 2440 TCCATCTCTGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2499
 DB 2303 TCTTCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2352
 QY 2500 TATCGGAGCGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2559
 DB 2363 TGTGTGTGCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2422
 QY 2560 TGAACAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2619
 DB 2423 TAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2482
 QY 2620 CTATTAAATGCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2679
 DB 2483 GTACTACATTTGTGCGGACCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 2542
 QY 2680 GAGTTTCTTGTGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2739
 DB 2543 AGCTACCTGCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2602
 QY 2740 TTACTACTCTAA 2752

DB 2603 TTATTATCTATTA 2615

RESULT 13

AAH14032 standard; cDNA; 1733 BP.

AAH14032;

26-JUN-2001 (first entry)

Human cDNA sequence SEQ ID NO:11143.

Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-0116126.

29-JUL-1999; 99UP-0248036.

27-AUG-1999; 99UP-0300253.

11-JAN-2000; 2000JP-0118767.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

Claim 8; SEQ ID 11143; 2537pp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end complementary strand of a polynucleotide comprising a sequence complementary to a sequence and an oligonucleotide comprising a 3'-end sequence, where the oligonucleotide which comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any special methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH9246 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 1733 BP; 435 A; 408 C; 451 G; 439 T; 0 other;

Query Match 6.9%; Score 238.8; DB 22; Length 1733;
 Best Local Similarity 52.1%; Pred. No. 2.2e-49;
 Matches 609; Conservative 0; Mismatches 548; Indels 12; Gaps 3;

QY	1594	AGATTGGTCCAAAGAAAACAAGAGGTGCGACCAATTAATTTAGTGTTAAGCCACTAGATAA	CTGG	1653
Db	3	AAACTGGGTTTAAGGAAGTAACACAGAGACCCTTCATCTTGCATATCCCCTGCACTTTCTGG	62	
QY	1654	GCTGTTGATCTATACCGCNAGNAATATTGAAGCAGCCAAATTCATTGATACAAAACTCATTT	1713	
Db	63	GGCACATTTTTTATCCCAAAGAGAGAANAATGGACCGAGCTCGAGAACCTGGTCAACATGTTGGA	122	
QY	1714	TAAAGTTACACCAAGCCATGGGCGATGCCAAATAGAAAAAGCAATAANTGNTTGAAGT---	1770	
Db	123	GAAGATAGCGCGCCCCATTTGGCATTCGCGTATAGCCCAACCGGCTGGTGTGAACATAAAGGA	182	
QY	1771	TGACAGAACTGAAGCGCTACTTAAGAGTCTTACAGCAAAAGGTCACAGCAG-----ACAC	1824	
Db	183	TGACCGAATPAGAGACTTATGTGAGAACCATTCAAATCCAGCTTAGGAGCTGAGGGGAAGAT	242	
QY	1825	CCAGATAGTTGTCTGCTGTTGTCAAGTAATCCGAAGGCAAAATACGATGCTATTAAAAA	1884	
Db	243	ACAGATGGTTGTTTGCATCATCATGGGCCCACTGATGATCTCTATATGGGGCCATCAATAA	302	
QY	1885	ATACCTGTGTACAGATGCCCCATCCCAAGTCAAGTGTGTGTGTGGCCCGAACTTAGSCAA	1944	
Db	303	GCTGTGCTGTGTGCAGTCCCCCAGTGCCCTCCCAAGTTGTCAATGTTTGAACCATTGGTCA	362	
QY	1945	ACAGCAAACTGTCAATGGCCATGCTCAAAAGATTGCCCTACAGATGAATGCGAAGATGGG	2004	
Db	363	GCCCAACAGGCTTCGGAGTGTGCCCAGAAAGATTTTACTTCAGATTAACGTGTAATTTGGG	422	
QY	2005	AGGAGAGCTCTGGAAGGTGGACATCCCCCTGAAGCTCGTGATGATCGTTGGCATCCATTG	2064	
Db	423	TGGTGAGCTCTGGGGAGTGATATTCCTCTGAAAACAGTTAATGCTGATCGSGATGATGCT	482	
QY	2065	TTACCATGACATGACACTCGGGCGGAGGTCAATCCAGGATTTGTTGCCAGCATCAATGA	2124	
Db	483	TTACCATGACCCAGTAGAGGATCGCTCCGTGGTTGGCTTGGTGGCAAGCATCAATCT	542	
QY	2125	AGGATGACCCGCTGGTTCTCAAGCTGCAATATTTCCAGATAGAGGACAGGAGCTGGTGA	2184	
Db	543	CACCTTCACAAAATGGTATTTCCCGGTGGTGTTCAGATGCCGATCNGAGATTTGTGA	602	
QY	2185	TGGGCTCAAAAGTCTGCTCGACAGCGGCTCTGAGGCTTGGAAATAGCTGCAATGATACAT	2244	
Db	603	CAGCTCGAGCTATGCTCTGTGGGTCTCTTAAAAAGTTTATAGAGTGAACCACTGTCT	662	
QY	2245	GCCACGCCGATCATCGTTACCGGATGGGCTAGGAGACGCCAGCTCAAAACACTGGT	2304	
Db	663	ACCAGAGAAGATTTGGTGTACCGTGATGSAAGTCTCTGATGGCCAACCTGAGACAGATTGC	722	
QY	2305	GAATACGAGTGCACAGATTTTTCGATTTGTCTAAATTCATTTGGTGTAGAGTTACAAACC	2364	
Db	723	CAACTATGAGATTCCTCAACTACAGAAGTGTTTTGAAGCTTTTGTAGA---ATTATCAGCC	779	
QY	2365	TAGNTAAOCGTAATTTGGTGGAGAAAAGAGTGAAACACAGATTTTTTGTCTCAGTCTGG	2424	
Db	780	CAAGATGGTGGTGTGTAGTTTCAGAGAGAAAATCACTACTATCTATCTATCTGCGTGTCC	839	
QY	2425	ASGAAGACTTCAGAAATCCACTTCCTGGNACAGTTATTGATGTAGAGGTTACAGACAGA	2484	
Db	840	TCAGAACTTTGTAACTCCCACTCTCGTGAAGCTGTGGTAGATCATCAATAAACAGCTGTGA	899	
QY	2485	ATGGTATGACTTTTTTATCTGTGAGCCAGGCTGTGAGAAGTGTGTAGTGTTCCTCCACACA	2544	
Db	900	GTGGGTGGATTTCTATCTTCTTGCCCATCATGTACCGCAGGCGCTGTGCATTCCTCGCA	959	
QY	2545	TTACAAATGTCACTATGACACACAGCGGCTTGAAGCCAGACCATACATACAGCGCTTTGACCTA	2604	
Db	960	TTATGTCTGTGTTCTCAACACCCGAAACCTTGAGCCCTGTATCATATGACAGAGCGCTGACTTT	1019	
QY	2605	CAAGCTGTGCCACATCTATTACAACCTGGCCAGGTGTCAATTCGTGTTCTGCTCTCTTGGCA	2664	
Db	1020	CAAACTGTGCCACATGACTGGAATTTGSCCTGGCACCATCAGAGTTTCAGCTCTCTTGCAA	1079	
QY	2665	GTACGCCCAACAGCTGGCTTTCTTTGTGGCCAGATTTTACAGAGAGGCAATCTGCTC	2724	

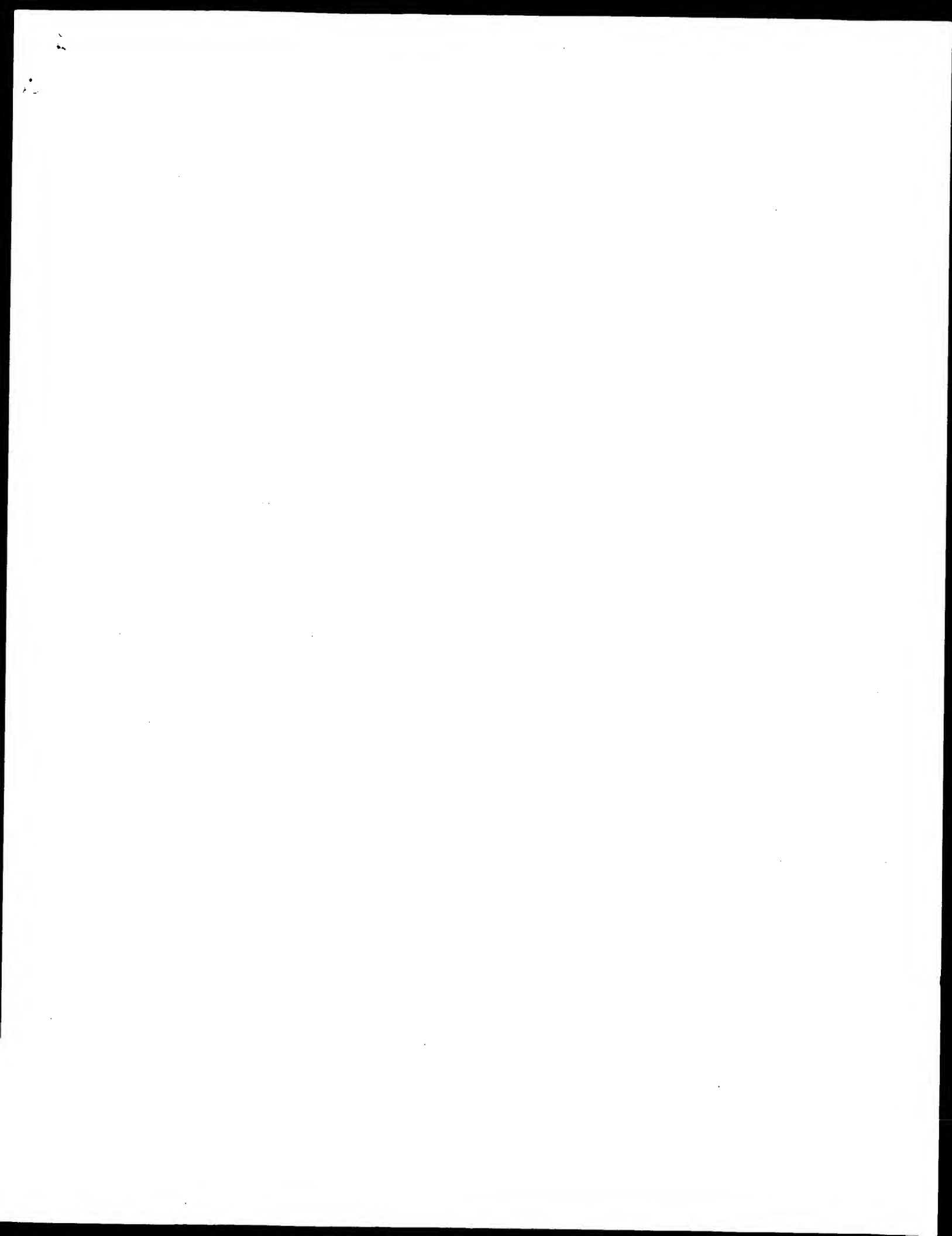
Db	1080	GTATGCCCAACAGCTAGCTTTCTGTGAGGACACATCTTGGCATCATGACAGCCATCCA	1139
Qy	2725	ACTGTCAAACCGCTTTACTACCTCTAAC	2753
Db	1140	GCTGTGCGAGAACCTGTCTCTCTGTGAC	1168
RESULT 14			
ABL10570			
ID	ABL10570	standard; cDNA; 5671 BP.	
XX			
AC	ABL10570;		
XX			
DT	26-MAR-2002	(first entry)	
XX			
DE	Drosophila melanogaster	expressed polynucleotide SEQ ID NO 26192.	
XX			
KW	Drosophila;	developmental biology; cell signalling; insecticide;	
KW	pharmaceutical; gene; ss.		
XX			
OS	Drosophila melanogaster.		
XX			
PN	WO200171042-A2.		
XX			
PD	27-SEP-2001.		
XX			
PF	23-MAR-2001;	2001WO-US09231.	
XX			
PR	23-MAR-2000;	2000US-191637P.	
PR	11-JUL-2000;	2000US-0614150.	
XX			
PA	(PEKE)	PE CORP NY.	
XX			
PI	Venter JC, Adams M, Li PWD, Myers EW;		
XX			
DR	WPI; 2001-655860/75.		
DR	P-PSDB; ABB66467.		
XX			
PT	New isolated nucleic acid detection reagent for detecting 1000 or more		
PT	genes from Drosophila and for elucidating cell signalling and cell-cell		
PT	interactions		
XX			
PS	Claim 1; SEQ ID NO 26192; 2lpp + Sequence Listing; English.		
XX			
CC	The invention relates to an isolated nucleic acid detection reagent		
CC	capable of detecting 1000 or more genes from Drosophila. The invention		
CC	useful in developmental biology and in elucidating cell signalling and		
CC	cell-cell interactions in higher eukaryotes for the development of		
CC	insecticides, therapeutics and pharmaceutical drugs. The invention		
CC	discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA		
CC	sequences (ABL01840-ABL16175) and the encoded proteins		
CC	(ABBS7737-ABBY2072).		
CC	The sequence data for this patent did not form part of the printed		
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp.wipo.int/pub/published_pct_sequences.		
XX			
SQ	Sequence 5671 BP; 1723 A; 1127 C; 1131 G; 1690 T; 0 other;		
Query Match			
Best Local Similarity 52.8%; Score 184.6; DB 23; Length 5671;			
Matches 422; Conservative 0; Mismatches 375; Indels 3; Gaps			
Qy	1956	TCATGGCAATGTACAAAGATTGCCCTACAGATCAATGCGAGATGGAGAGCTCT	2015
Db	3759	TGATGTGATCGCCACAAAGGTGGTTATTCAAATGAACGCCAAATTGATGGAGCTCCCT	3818
Qy	2016	GSAGGGTGCACATCCCCCTCGAAGCTCGTGATGATCGTTGGCATCGATTGTACCATGACA	2075
Db	3819	GGCAGGTAGTGTATCCCCCTCCACGGTCTGATGACGTGTGGTTCCGATCTCGCATTCAC	3878
Qy	2076	TGACAGCTGGCGAGGTCAATCGCAGGATTTGTTCACGAT---CAATGAAGGATGA	2132

Db 3879 CGAAGATTAAGAACAAATCATATGGGGCTTTCCTTGCACCAACGAGACCAAGAGAGCTTT 3938
 Qy 2133 CCGGCTGGTCTCAGGCTGCATATTTTCAGATAGAGGACAGAGAGCTGTAGATGGGCTCA 2192
 Db 3939 TCCGGCTACTTCTCCACCTTAAACGACATTTAAGGGGCAAGAGTTCCTCCAGAGAGATGT 3998
 Qy 2193 AAGTCTGCTGACAGCGGCTGTAGAGGCTTGGATAGCTGCATAGAGTAACTGCCCCAGCC 2252
 Db 3999 CCGTGAACATGGCGTGGCCCTGAGGTGTATCAGAGCAACACCGTTCCTTGGCAGAGC 4058
 Qy 2253 GGAATCTCTGTACCGGCTAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2312
 Db 4059 GCAATCTCTCTTCTTCCGACGAGGTGTGGGAGTGTGCTACCTACAGGTGGTAAACAGCG 4118
 Qy 2313 AAGTGCACAGTGTGTGATTTGATTTAAATCCATTTGATGAGGTTACACCTTGAGTAA 2372
 Db 4119 AGGTGAACACCTTAAAGGACAGGCTGACGAAATTTACAAATCAGCTGGCAACAGAGG 4178
 Qy 2373 CGGTAAATTTGTGTAAGAAAGAGTGAACACAGATTTTTCCTCAGTGTGAGGAGAGAC 2432
 Db 4179 GCTGTGCAATGACATTTATTAATTTGATTCAGGCGCATTAATTTCTGCTACTTTACTGGGC 4238
 Qy 2433 TTGAGATCCACTTCTCTGGAACAGTTATGATGTAGAGTTACAGACCAAGATGTATG 2492
 Db 4239 ATGCAACCCAGTTCCGGGCACTGTAGTGTAGTATTTACTTGTCCAGAGGCTTACG 4298
 Qy 2493 ACTTTTATTCGTGACCGAGGCTGTGAGAGTGTATGTTTCTCCACATTTACATG 2552
 Db 4299 ACTTCTCTAGTGTCCAGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4358
 Qy 2553 TCAATCTATGACACAGCGGCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2612
 Db 4359 TTAATTTCTGCAACATGAGGAGTGAACCGGATTAAGTGTAGTGTATTTCTCTTAAGATGA 4418
 Qy 2613 GCCACATCTATTAACAAGTGGCCAGGTGTATGTTGTTCTCTCTTCCGATAGCCCC 2672
 Db 4419 CCAATATGTTCTCAATTAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4478
 Qy 2673 ACAAGCTGCTTTTCTTGTGGCCAGAGTATTCACAGAGAGGCAATCTGTACTGTCAA 2732
 Db 4479 ACAATTTGGCTTCTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4538
 Qy 2733 ACCGCTTTACTACTCTTAA 2752
 Db 4539 ATCAATTTACTTTTGTAA 4558

RESULT 15
 ABN62610
 ID ABN62610 standard; cDNA; 548 BP.
 AC ABN62610;
 AC 28-JUN-2002 (first entry)
 DE Human cancer related polynucleotide SEQ ID NO 2577.
 DE Human, cytochrome c, gene expression; gene mapping; tissue profiling;
 DE gene therapy; cancer; tumour; gene; ss.
 OS Homo sapiens.
 PN WO200214500-A2.
 PD 21-FEB-2002.
 PF 16-AUG-2001; 2001WO-US25840.
 PR 16-AUG-2000; 2000US-226326P.
 PA (CHIR) CHIRON CORP.
 PA (HISE-) HISEQ INC.
 XX

P1 Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;
 P2 Lanson G, Scott EM, Zhang G, Kaesam A, Pot D, Labat I;
 P3 WIPO: 2002-241905/29.
 P4 New nucleic acid for producing a polypeptide, detecting differentially
 P5 expressed genes correlated with a cancerous state of a mammalian cell,
 P6 and inhibiting tumor growth -
 P7
 P8 Claim 1: SEQ ID NO 2577; 883bp + Sequence Listing; English.
 P9
 P10 The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
 P11 with cytostatic activity. The polynucleotide is used to produce a
 P12 polypeptide, to detect differentially expressed genes correlated with a
 P13 cancerous state of a mammalian cell and to inhibit tumour growth. The
 P14 polynucleotide is used as a probe in mapping and tissue profiling. The
 P15 encoded polypeptide and antibodies to the polypeptide can also be used
 P16 for therapeutic and diagnostic purposes. The polynucleotide is useful for
 P17 gene therapy.
 P18 Note: The sequence data for this patent did not form part of the printed
 P19 specification, but was obtained in electronic format directly from WIPO
 P20 at ftp.wipo.int/pub/published_pct_sequences.
 P21
 P22 Sequence 548 BP; 151 A; 116 C; 146 G; 134 T; 1 other;
 P23
 P24 Query Match 4.9%; Score 170.8; DB 24; Length 548;
 P25 Best Local Similarity 57.0%; Pred. No. 1,4e-32;
 P26 Matches 310; Conservative 0; Mismatches 234; Indels 0; Gaps 0;
 P27
 P28 Qy 1943 AACAAGCAACTGTATGATGCAATGCTTCAAGATGTCCTTCAAGATGATGCAAGATG 2002
 P29 Db 5 AACAAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 64
 P30 Qy 2003 GAGAGAGAGCTGTGAGAGGATGAGATCCCTTGAAGTGTGATGATGATGATGATGATGATG 2062
 P31 Db 65 GAGAGGAGAGCTGTGAGAGGATGAGATCCCTTGAAGTGTGATGATGATGATGATGATGATG 124
 P32 Qy 2063 TGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2122
 P33 Db 125 GTCTGTAAAGATGCACTGACAGAGGAGGATGATGATGATGATGATGATGATGATGATGATG 184
 P34 Qy 2123 GAGGAGATGACCGGCTGTGCTTCAAGTGTATTTTCAAGTGTATGAGAGGAGGAGGAGGAGG 2182
 P35 Db 185 CCCAGATTCACAGGAGTGTGCTTCCGCTGTATCTTCAAGAGAGATGATGATGATGATGATG 244
 P36 Qy 2183 GATGGGCTCAAGTGTGCTGCAAGGCTGTGAGGCTTGGAAATAGCTGCAATGATGATGATG 2242
 P37 Db 245 GATGGCTTGAAGTGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 304
 P38 Qy 2243 ATGCCAGCGGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2302
 P39 Db 305 TTGCCAGACAGGATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 364
 P40 Qy 2303 GTGAATACAGAGTGTGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2362
 P41 Db 365 ATTGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 424
 P42 Qy 2363 CTGAGATTAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2422
 P43 Db 425 TCAAGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 484
 P44 Qy 2423 GAGAGAGAGCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2482
 P45 Db 485 AACGCACTGTACAGAGCCCACTTGGACATGTTGTGATGATGATGATGATGATGATGATGATG 544
 P46 Qy 2483 GAAT 2486
 P47 Db 545 GAAT 548

Search completed: January 9, 2003, 19:04:28
 Job time : 504 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2003, 18:55:10 ; Search time 88 Seconds
(without alignments)
12099.792 Million cell updates/sec

Title: US-09-873-737A-5

Perfect score: 3472

Sequence: 1 gggcggaattggggccgagc.....taaataaatgatgtcc 3472

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/pdata/1/ina/5A_COMB.seq:*
2: /cgn2_6/pdata/1/ina/5B_COMB.seq:*
3: /cgn2_6/pdata/1/ina/6A_COMB.seq:*
4: /cgn2_6/pdata/1/ina/6B_COMB.seq:*
5: /cgn2_6/pdata/1/ina/PCUS_COMB.seq:*
6: /cgn2_6/pdata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	60.8	1.8	2365	4	US-09-249-697A-5	Sequence 5, Appl
2	60.8	1.8	2365	4	US-09-249-697A-18	Sequence 18, Appl
3	60.8	1.8	2365	4	US-09-363-316B-5	Sequence 5, Appl
4	60.8	1.8	2365	4	US-09-363-316B-23	Sequence 23, Appl
5	57.6	1.7	7218	1	US-08-232-463-14	Sequence 14, Appl
6	57	1.6	1613	3	US-08-838-830-28	Sequence 28, Appl
7	52.8	1.5	2217	4	US-09-244-314-1	Sequence 1, Appl
8	52.8	1.5	2217	4	US-09-498-959-1	Sequence 1, Appl
9	52.4	1.5	7218	1	US-08-232-463-14	Sequence 14, Appl
10	51.8	1.5	19124	2	US-08-487-826B-13	Sequence 13, Appl
11	51	1.5	797	1	US-08-592-126-73	Sequence 73, Appl
12	50.4	1.5	3388	4	US-09-141-206-1	Sequence 1, Appl
13	50.2	1.4	4203	2	US-08-866-757-1	Sequence 1, Appl
14	50.2	1.4	4203	4	US-09-153-593-1	Sequence 1, Appl
15	50	1.4	658	4	US-08-998-416-595	Sequence 595, App
16	49.4	1.4	5852	1	US-07-867-106-2	Sequence 2, Appl
17	49.4	1.4	1230	4	US-09-387-574-5	Sequence 5, Appl
18	49.2	1.4	1230	4	US-09-668-096-5	Sequence 6, Appl
19	49.2	1.4	2447	2	US-08-628-417-6	Sequence 14, Appl
20	48.8	1.4	2447	2	US-09-014-969-14	Sequence 1, Appl
21	47	1.4	3396	4	US-08-989-299-1	Sequence 11, Appl
22	45.6	1.3	2251	4	US-08-991-677-11	Sequence 1, Appl
23	45.6	1.3	6243	2	US-09-056-075-1	Sequence 3, Appl
24	45.4	1.3	6124	4	US-08-213-419B-3	Sequence 51, Appl
25	44.8	1.3	163	4	US-09-182-145-52	Sequence 52, Appl
26	44.8	1.3	163	4	US-09-182-145-52	Sequence 52, Appl
27	44.8	1.3	488	4	US-08-991-789A-180	Sequence 180, App

28	44.8	1.3	488	4	US-09-062-451-180	Sequence 180, App
29	44.8	1.3	488	4	US-09-598-326-180	Sequence 180, App
30	44.8	1.3	507	4	US-08-991-789A-182	Sequence 182, App
31	44.8	1.3	507	4	US-09-062-451-182	Sequence 182, App
32	44.8	1.3	507	4	US-09-598-326-182	Sequence 83, Appl
33	44.8	1.3	1222	4	US-08-861-774E-83	Sequence 83, Appl
34	44.8	1.3	2064	4	US-09-178-109-3	Sequence 1, Appl
35	44.8	1.3	2121	4	US-09-178-109-3	Sequence 93, Appl
36	44.8	1.3	5506	4	US-09-004-838-93	Sequence 5, Appl
37	44	1.3	44	1	US-08-381-572-5	Sequence 6, Appl
38	44	1.3	44	1	US-08-592-820-5	Sequence 6, Appl
39	44	1.3	44	1	US-08-592-820-6	Sequence 13, Appl
40	44	1.3	44	2	US-08-670-707A-13	Sequence 13, Appl
41	44	1.3	44	4	US-09-037-601-13	Sequence 5, Appl
42	44	1.3	44	4	US-09-234-393-5	Sequence 19, Appl
43	44	1.3	44	4	US-09-124-238A-19	Sequence 23, Appl
44	44	1.3	44	4	US-09-323-873A-23	
45	44	1.3	44	4		

ALIGNMENTS

RESULT 1
US-09-249-697A-5
Sequence 5, Application US/09249697A
Patent No. 6392018
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Yeung, George
TITLE OF INVENTION: NOVEL EGF MOTIF PROTEIN OBTAINED FROM A CDNA LIBRARY OF FETAL LIVER SPLEEN
FILE REFERENCE: 24011-727
CURRENT APPLICATION NUMBER: US/09/249,697A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 08/968,800
PRIOR FILING DATE: 1997-11-22
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 5
LENGTH: 2365
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (205)...(1866)
US-09-249-697A-5

Query Match 1.8%; Score 60.8; DB 4; Length 2365;
Best Local Similarity 90.3%; Pred. No. 9.7e-07;
Matches 65; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 58 GATTCATCTAATACGACTACCTATAGGCTCGAGCGCCGCCGACAGTGTGAGGT 117
DB 7 GATTCATCTAATACGACTACCTATAGGCTCGAGCGCCGCCGACAGTGTGAGGT 66

QY 118 GCAAGACCCAGG 129
DB 67 GACAGACCCCG 78

RESULT 2
US-09-249-697A-18
Sequence 18, Application US/09249697A
Patent No. 6392018
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Yeung, George
TITLE OF INVENTION: NOVEL EGF MOTIF PROTEIN OBTAINED FROM A CDNA LIBRARY OF FETAL LIVER SPLEEN
FILE REFERENCE: 24011-727
CURRENT APPLICATION NUMBER: US/09/249,697A
CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 08/968,800
; PRIOR FILING DATE: 1997-11-22
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 2365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (205)..(1863)
US-09-249-697A-18

Query Match 1.8%; Score 60.8; DB 4; Length 2365;
Best Local Similarity 90.3%; Pred. No. 9.7e-07;
Matches 65; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 58 GATTCATCCTAATACGACTCACTATAGGCTCGAGCGCGCCGGCGAGGTCTGAGGT 117
DB 7 GATTCATCCTAATACGACTCACTATAGGCTCGAGCGCGCCGGCGAGGTCTGAGG 66

QY 118 GCAAGGACCAGG 129
DB 67 GACAGCACCCGG 78

RESULT 3

US-09-363-316B-5
; Sequence 5, Application US/09363316B
; Patent No. 6392019
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; TITLE OF INVENTION: EGF MOTIF PROTEIN MATERIALS AND METHODS
; FILE REFERENCE: 28110/35852
; CURRENT APPLICATION NUMBER: US/09/363,316B
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: US 09/249,697
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 08/968,800
; PRIOR FILING DATE: 1997-11-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 2365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (205)..(1866)
; NAME/KEY: misc feature
; LOCATION: (357)
; OTHER INFORMATION: Xaa = any amino acid
US-09-363-316B-5

Query Match 1.8%; Score 60.8; DB 4; Length 2365;
Best Local Similarity 90.3%; Pred. No. 9.7e-07;
Matches 65; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 58 GATTCATCCTAATACGACTCACTATAGGCTCGAGCGCGCCGGCGAGGTCTGAGGT 117
DB 7 GATTCATCCTAATACGACTCACTATAGGCTCGAGCGCGCCGGCGAGGTCTGAGG 66

QY 118 GCAAGGACCAGG 129
DB 67 GACAGCACCCGG 78

RESULT 4

US-09-363-316B-23
; Sequence 23, Application US/09363316B
; Patent No. 6392019
; GENERAL INFORMATION:

; APPLICANT: Ford, John
; APPLICANT: Yeung, George
; TITLE OF INVENTION: EGF MOTIF PROTEIN MATERIALS AND METHODS
; FILE REFERENCE: 28110/35852
; CURRENT APPLICATION NUMBER: US/09/363,316B
; CURRENT FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: US 09/249,697
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 08/968,800
; PRIOR FILING DATE: 1997-11-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 2365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (205)..(1863)
US-09-363-316B-23

Query Match 1.8%; Score 60.8; DB 4; Length 2365;
Best Local Similarity 90.3%; Pred. No. 9.7e-07;
Matches 65; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 58 GATTCATCCTAATACGACTCACTATAGGCTCGAGCGCGCCGGCGAGGTCTGAGGT 117
DB 7 GATTCATCCTAATACGACTCACTATAGGCTCGAGCGCGCCGGCGAGGTCTGAGG 66

QY 118 GCAAGGACCAGG 129
DB 67 GACAGCACCCGG 78

RESULT 5

US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149


```

RESULT 8
US-09-498-959-1
; Sequence 1, Application US/09498959
; Patent No. 6410240
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; APPLICANT: Yowe, David
; TITLE OF INVENTION: RGS-Containing Molecules and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 5800-129
; CURRENT APPLICATION NUMBER: US/09/498,959

```


INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:
 LENGTH: 19124 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-487-826B-13

Query Match

1.5%; Score 51.8; DB 2; Length 19124;
 Best Local Similarity 51.0%; Pred. No. 0.0011;

Matches 175; Conservative 0; Mismatches 162; Indels 6; Gaps 2;

QY 2776 TTTTCTTTTGAAGACTTTGGGATTTTAAAGCTTTTAACTTTTAACTG 2835
 DB 15720 TTTTCTTTTGAAGACTTTGGGATTTTAAAGCTTTTAACTG 15661
 QY 2836 TTAATCTTCGATGAACTGGGAAAGGATTAAGAGATCTAGCATTTATTTCTAGCA 2895
 DB 15660 TTTCTTTGTTTATTTTATTTTATATCAATTTTATATTAATAATTTTAAATTT 15601
 QY 2896 TTGCTAT--TCACGGCTCTCTATTTATATGTAATAATTAAGATTTATTTATCT 2953
 DB 15600 TTTTGTGATTAATCTTTTCAATTTTATTTATCAAAATTTATATTTATATATTTT 15541
 QY 2954 TCTGTTTCTCATAGATTTTGTGAGCATTTTGTATTTTGAAGAAATGTGATA 3013
 DB 15540 TATATTTTATTAATAATTTTCTCTTTTATTTTATTTTATTTTATTTTATTTT 15485
 QY 3014 AGATACTTGGTATTAACAAGACTCTGAGAGATTTGAAAGTGTGAGATTTA 3073
 DB 15484 AATTTTATTAATTTTCAATTTTCTTTTATTTTAAATTTGTTTATATTTTC 15425
 QY 3074 CTTAAGTACTTTCAGAGTGAGCAAGTCTTATTAATTAAC 3116
 DB 15424 TTTTATTAATTAATCAATATATTAATAATTAATTAATAAC 15382

RESULT 11

US-08-592-126-73/c

; Sequence 73; Application US/08592126

; Patent No. 5821091

; GENERAL INFORMATION:

; APPLICANT: Gregory Dolganov

; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory

; TITLE OF INVENTION: Polypeptides

; NUMBER OF SEQUENCES: 151

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dehlinger & Associates

; STREET: 350 Cambridge Avenue, Suite 250

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/592.126

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Shultz, Charles K.

; REGISTRATION NUMBER: 38,615

; REFERENCE/DOCKET NUMBER: 4600-0111

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 324-0880

; TELEFAX: (415) 324-0960

; INFORMATION FOR SEQ ID NO: 73:

SEQUENCE CHARACTERISTICS:

LENGTH: 797 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: G248b.seq
 US-08-592-126-73

Query Match

1.5%; Score 51; DB 1; Length 797;
 Best Local Similarity 91.5%; Pred. No. 0.00027;

Matches 54; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 51 GGAATTCATTCCTCATTAATGACCTCATTAGGGCTGAGCGCGCCGCGGACAG 109
 DB 797 GAATTCGCTTCATTCCTCATTAATGACCTCATTAGGGCTGAGCGCGCCGCGGACAG 739

RESULT 12

US-09-141-206-1

; Sequence 1; Application US/09141206

; Patent No. 6187559

; GENERAL INFORMATION:

; APPLICANT: Steed, Paul M.

; TITLE OF INVENTION: Amino Acid Sequence of Human PLD2A

; FILE REFERENCE: 4-30148/P1/CGC1954/R

; CURRENT APPLICATION NUMBER: US/09/141,206

; CURRENT FILING DATE: 1998-08-27

; EARLIER APPLICATION NUMBER: 60/057,802

; EARLIER FILING DATE: 1997-08-28

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 1

; LENGTH: 3388

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (162)..(2963)

; OTHER INFORMATION: Human PLD2

; US-09-141-206-1

Query Match

1.5%; Score 50.4; DB 4; Length 3388;
 Best Local Similarity 98.1%; Pred. No. 0.00096;

Matches 51; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 62 CCATCTTAATGACACTCATATAGGGCTGAGCGCGCGCGGACAGGCTG 113
 DB 1 CCATCTTAATGACACTCATATAGGGCTGAGCGCGCGCGGACAGGCTG 52

RESULT 13

US-08-866-757-1

; Sequence 1; Application US/08866757

; Patent No. 5858716

; GENERAL INFORMATION:

; APPLICANT: ELISHOURBAGY, NABIL A

; APPLICANT: LI, XIATONG

; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RATNER & PRESTIA

; STREET: P.O. BOX 980

; CITY: VALLEY FORGE

; STATE: PA

; COUNTRY: USA

; ZIP: 19482

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/866,757
FILING DATE: 30-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70055
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4203 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-866-757-1

Query Match
Best Local Similarity 1.4%; Score 50.2; DB 2; Length 4203;
Matches 70; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 58 GATTCATCCCTAATACGACTACATAGGCTCGAGCGCGCGCGCGCGGCGAGGTCTGAGGT 117
Db 1 GCTTCATCCCTAATACACTCATTATAGGCTCGAGCGCGCGCGCGCGGCGAGGTCTGAC 60

QY 118 GCAAGGACCGAGGACTAGGCGGCGGCGGCGGCTCCCAAGAAATAG 160
Db 61 GGAGGTGCTGTGACCCCTCAGCAATCTGCCACCCCTACAG 103

RESULT 14
US-09-153-593-1
Sequence 1, Application US/09153593A
Patent No. 6174994
GENERAL INFORMATION:
APPLICANT: ELSHOURBAGY, NABIL A
APPLICANT: LI, XIAOTONG
APPLICANT: BERGSWA, DERK J
TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
FILE REFERENCE: GH-70055-1
CURRENT APPLICATION NUMBER: US/09/153,593A
CURRENT FILING DATE: 1998-09-15
EARLIER APPLICATION NUMBER: 08/866,757
EARLIER FILING DATE: 1997-05-30
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 4203
TYPE: DNA
ORGANISM: HOMO SAPIENS
US-09-153-593-1

Query Match
Best Local Similarity 1.4%; Score 50.2; DB 4; Length 4203;
Matches 70; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 58 GATTCATCCCTAATACGACTACATAGGCTCGAGCGCGCGCGCGGCGAGGTCTGAGGT 117
Db 1 GCTTCATCCCTAATACACTCATTATAGGCTCGAGCGCGCGCGCGGCGAGGTCTGAC 60

QY 118 GCAAGGACCGAGGACTAGGCGGCGGCGGCGGCTCCCAAGAAATAG 160

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 595:
SEQUENCE CHARACTERISTICS:
LENGTH: 658 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1408RP
US-08-998-416-595

Query Match
Best Local Similarity 1.4%; Score 50; DB 4; Length 658;
Matches 151; Conservative 0; Mismatches 150; Indels 1; Gaps 1;

QY 2776 TTTTCTTTTGAATGACTTTGGGATTTTTTTAAGCTTTTATTACTTTTCTTAACTG 2835
Db 275 TATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 334

QY 2836 TTATCTTTCTGGATGAACCTTGGGAAGGATAGGAGATCTAGCATTTTATTCTAGCA 2895
Db 335 ATTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 394

QY 2896 TTGCTATTTCACCGCTTCTTATTATTATATGTAATAAATTAAGATTATTATTCCTC 2955
Db 395 TTTTATTATTATGTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTA 453

QY 2956 TTGTTTCTCATAGATATTTTGTGAGCATTTTTTTTGTGAGCAATTTTTTTTGAAGAAATG 3015
```


Db 1 GGGCGAATTGGCCCGGACGTCGATGCTCCGCGCGCCATGGCGCGCGGGAATTCGAT 60
 Qy 61 TCCATCTTAATACGACTCACTATAGGCTCGAGCGGCCCGCGGCGAGGTCTGAGGTGA 120
 Db 61 TCCATCTTAATACGACTCACTATAGGCTCGAGCGGCCCGCGGCGAGGTCTGAGGTGA 120
 Qy 121 AGGACGAGGACTAGGCGGCGGCGGCTCCAGGAATAGAAAATAGATCTGGGAGAGC 180
 Db 121 AGGACGAGGACTAGGCGGCGGCGGCTCCAGGAATAGAAAATAGATCTGGGAGAGC 180
 Qy 181 CCGAGCCAGAGCCAGAGGAGGCGCGGCTCAGGAGACAGCGCAGCTGGTGGGCTCCAC 240
 Db 181 CCGAGCCAGAGCCAGAGGAGGCGCGGCTCAGGAGACAGCGCAGCTGGTGGGCTCCAC 240
 Qy 241 TGCAGCTCAGCACTGGTGTATATTCAGGCTAGGCTCAGCGGCGCAACAGAGGAGGGA 300
 Db 241 TGCAGCTCAGCACTGGTGTATATTCAGGCTAGGCTCAGCGGCGCAACAGAGGAGGGA 300
 Qy 301 ATTAATTTGGCCGTGGACGCGCAGAGGAACAGCAGGAGGAACAGCAGGAGTCAAGGACT 360
 Db 301 ATTAATTTGGCCGTGGACGCGCAGAGGAACAGCAGGAGGAACAGCAGGAGTCAAGGACT 360
 Qy 361 CCAGATATCTGCTGGATTTTCAGAGGTATTCGNTAGCAGAGAGAGGTCTGCTGAGAGA 420
 Db 361 CCAGATATCTGCTGGATTTTCAGAGGTATTCGNTAGCAGAGAGAGGTCTGCTGAGAGA 420
 Qy 421 TTTTCATGATCTTTGGTGTGAATACAGGCGAGAACTAGACCATCTTAAAGNATCAAAAAC 480
 Db 421 TTTTCATGATCTTTGGTGTGAATACAGGCGAGAACTAGACCATCTTAAAGNATCAAAAAC 480
 Qy 481 AGGTTCTTCAGGCNTATAGTAAGGTTAAGCACTAACCATTTCCGGCTGACATCCCGTCC 540
 Db 481 AGGTTCTTCAGGCNTATAGTAAGGTTAAGCACTAACCATTTCCGGCTGACATCCCGTCC 540
 Qy 541 CCAGTGGGCTTATATCAGTACTCATTTAGACTATAACCCACTGATGGAAGCCAGAGACT 600
 Db 541 CCAGTGGGCTTATATCAGTACTCATTTAGACTATAACCCACTGATGGAAGCCAGAGACT 600
 Qy 601 CCGTTACGCTCTTTTTCACGAGACTTAATTTGGAAGTGTCTATCTTTTTCATGCG 660
 Db 601 CCGTTACGCTCTTTTTCACGAGACTTAATTTGGAAGTGTCTATCTTTTTCATGCG 660
 Qy 661 AACGATATTTTTCCTTAAGACTACAGCAAGAGTTTACTGAAAGTTTCTAGTAAGAC 720
 Db 661 AACGATATTTTTCCTTAAGACTACAGCAAGAGTTTACTGAAAGTTTCTAGTAAGAC 720
 Qy 721 CCGAATGGAGAGGATGTAGGATAACGATCACTTTTAAAGTAACTTCCACCTACATC 780
 Db 721 CCGAATGGAGAGGATGTAGGATAACGATCACTTTTAAAGTAACTTCCACCTACATC 780
 Qy 781 ACCAATCTGTTTTCAGTCTCTATATATTTTTCAGAGGCTTTTGAAGTCAATGAAATTT 840
 Db 781 ACCAATCTGTTTTCAGTCTCTATATATTTTTCAGAGGCTTTTGAAGTCAATGAAATTT 840
 Qy 841 GCAACAAATTTGAGCAATTTTATTAACCAATGATGATGATTTTTCAGGTCACAG 900
 Db 841 GCAACAAATTTGAGCAATTTTATTAACCAATGATGATGATTTTTCAGGTCACAG 900
 Qy 901 GTTGGTGAATTTGGCTGGCTTCACTACTTCCATCTTCAGTATGAAAACAGCATCATGCT 960
 Db 901 GTTGGTGAATTTGGCTGGCTTCACTACTTCCATCTTCAGTATGAAAACAGCATCATGCT 960
 Qy 961 CTGCACCTGAGCTTTCAGCAATGCTTTCAGAGTCTGCTGAGTCTGCTGAGTCTGAGT 1020
 Db 961 CTGCACCTGAGCTTTCAGCAATGCTTTCAGAGTCTGCTGAGTCTGCTGAGTCTGAGT 1020
 Qy 1021 CTTTTATCATCAGACAGAGAAACATAAATTTTCAAGAACAAAGTTTCCAAAGAAATTAAGG 1080
 Db 1021 CTTTTATCATCAGACAGAGAAACATAAATTTTCAAGAACAAAGTTTCCAAAGAAATTAAGG 1080
 Qy 1081 TTTAGTGTGCTTTTACCAAGTATAACATTAAGACATACAGAGTGAATGATTTGACTGGGA 1140
 Db 1081 TTTAGTGTGCTTTTACCAAGTATAACATTAAGACATACAGAGTGAATGATTTGACTGGGA 1140

Qy 1141 CCGAATCCCAAGAGACACCTTTAAGAAAGCCGACCGCTCTGGGCTCAGCTTCTTAGAATA 1200
 Db 1141 CCGAATCCCAAGAGACACCTTTAAGAAAGCCGACCGCTCTGGGCTCAGCTTCTTAGAATA 1200
 Qy 1201 CTACAGGAAGCAATACAAAGAGATCACCGACTCTGAAGCAGCCTCTCTTGGTCAAGCA 1260
 Db 1201 CTACAGGAAGCAATACAAAGAGATCACCGACTCTGAAGCAGCCTCTCTTGGTCAAGCA 1260
 Qy 1261 GCGAAGAGAGAGCGCGGCGCTCTGGGCGGACACTGCGAGGCGCTGCCATGCTCTCTGA 1320
 Db 1261 GCGAAGAGAGAGCGCGGCGCTCTGGGCGGACACTGCGAGGCGCTGCCATGCTCTCTGA 1320
 Qy 1321 GCTCTGCTATCTTACAGGCTCTAACTGATATAAATCGTAATGATTTTAACTGATGAAGA 1380
 Db 1321 GCTCTGCTATCTTACAGGCTCTAACTGATATAAATCGTAATGATTTTAACTGATGAAGA 1380
 Qy 1381 CTTAGCCCTTTCATACAGACTTAACCTCCAGAGCAAGGCGAGCTTGAAGTGGAGCACTCAT 1440
 Db 1381 CTTAGCCCTTTCATACAGACTTAACCTCCAGAGCAAGGCGAGCTTGAAGTGGAGCACTCAT 1440
 Qy 1441 TGATTTACATTTCAATAAAGCAATATGTTCAAGGAGCTTTCGAGACTTGGGTTTGAAGTT 1500
 Db 1441 TGATTTACATTTCAATAAAGCAATATGTTCAAGGAGCTTTCGAGACTTGGGTTTGAAGTT 1500
 Qy 1501 TGATTTCAACTTTACTGCTCTTCTCAGGAAGAAATTTGCAAAACAGAAAGATTCACCAAG 1560
 Db 1501 TGATTTCAACTTTACTGCTCTTCTCAGGAAGAAATTTGCAAAACAGAAAGATTCACCAAG 1560
 Qy 1561 TGGAAAACATTTGATTAATCAATCCAAATTTGAGATTTGGTCCAAAGAAACAGAGGTC 1620
 Db 1561 TGGAAAACATTTGATTAATCAATCCAAATTTGAGATTTGGTCCAAAGAAACAGAGGTC 1620
 Qy 1621 ACCATTAATTTAGTGTAAAGCCACTAGATACTGGCTGTTGATATACCGGAAGAAATTA 1680
 Db 1621 ACCATTAATTTAGTGTAAAGCCACTAGATACTGGCTGTTGATATACCGGAAGAAATTA 1680
 Qy 1681 TGAAGCAGCAATTTCAATGATACAAAATCTATTTAAAGTTTACCCAGGCTTGGGATGCA 1740
 Db 1681 TGAAGCAGCAATTTCAATGATACAAAATCTATTTAAAGTTTACCCAGGCTTGGGATGCA 1740
 Qy 1741 AATGAGAAAAGCAATTAATGATTTCAAGTGTGATGACAACTTAAAGCTTCTTAAAGCTTT 1800
 Db 1741 AATGAGAAAAGCAATTAATGATTTCAAGTGTGATGACAACTTAAAGCTTCTTAAAGCTTT 1800
 Qy 1801 ACAGAAAAGGTCACAGCAGACACCCAGATAGTGTCTGTTGTCAAGTAAATCGGAA 1860
 Db 1801 ACAGAAAAGGTCACAGCAGACACCCAGATAGTGTCTGTTGTCAAGTAAATCGGAA 1860
 Qy 1861 GGCAAAATAGATGCTTATTTAAATAATCTGTGTACAGATTTGCCCTACCCCAAGTCAAGT 1920
 Db 1861 GGCAAAATAGATGCTTATTTAAATAATCTGTGTACAGATTTGCCCTACCCCAAGTCAAGT 1920
 Qy 1921 TGTGTGGCCGAACTTAGGCAACACAGAACTGTCTATGCGGCTTGTCTACAAAGATTGC 1980
 Db 1921 TGTGTGGCCGAACTTAGGCAACACAGAACTGTCTATGCGGCTTGTCTACAAAGATTGC 1980
 Qy 1981 CCTACAGATGAACTGCAAGATGGGAGAGACTCTGGAGGCTGGACATCCCTTGAAGCT 2040
 Db 1981 CCTACAGATGAACTGCAAGATGGGAGAGACTCTGGAGGCTGGACATCCCTTGAAGCT 2040
 Qy 2041 CGTGATGATCGTTGGCATCGATTTGATCCATGACATGACAGCTGGGCGGAGGTCATTCGC 2100
 Db 2041 CGTGATGATCGTTGGCATCGATTTGATCCATGACATGACAGCTGGGCGGAGGTCATTCGC 2100
 Qy 2101 AGGATTTGTTGCCAGCATCAATGAAGGATGACCCGCTGTTCTCAGCTGCAATTTCA 2160
 Db 2101 AGGATTTGTTGCCAGCATCAATGAAGGATGACCCGCTGTTCTCAGCTGCAATTTCA 2160
 Qy 2161 GGATAGAGACAGAGGCTGGTGTAGATGGCTCAAGTCTGCTCAAGCGGCTCTGAGGGC 2220
 Db 2161 GGATAGAGACAGAGGCTGGTGTAGATGGCTCAAGTCTGCTCAAGCGGCTCTGAGGGC 2220

QY 1721 ACACAGCCATGGCGATGCAATGAGAAAGCAATATGATTGAAGTGGATGACAGAACT 1780
Db 541 ACACAGCCATGGCGATGCAATGAGAAAGCAATATGATTGAAGTGGATGACAGAACT 600
QY 1781 GAAGCTTACTTAAAGAGTCTTACAGCAAAAGGTTCACAGCAGACACCCAGATAGTGTCTGT 1840
Db 601 GAAGCTTACTTAAAGAGTCTTACAGCAAAAGGTTCACAGCAGACACCCAGATAGTGTCTGT 660
QY 1841 GTGTGTCAAGTAAATCGGAAGGCAAAATACGATGCTATTAATAAATACCTGTGTACAGAT 1900
Db 661 GTGTGTCAAGTAAATCGGAAGGCAAAATACGATGCTATTAATAAATACCTGTGTACAGAT 720
QY 1901 TGCCCTACCCCAAGTCAAGTGTGTGTCGCCCAAGCTTAGGCAAAAGCAAACTGTGTATG 1960
Db 721 TGCCCTACCCCAAGTCAAGTGTGTGTCGCCCAAGCTTAGGCAAAAGCAAACTGTGTATG 780
QY 1961 GCCATTGCTACAAAGATTGCCCTACAGATGAATCGCAAGATGGGAGGAGCTCTGGAGG 2020
Db 781 GCCATTGCTACAAAGATTGCCCTACAGATGAATCGCAAGATGGGAGGAGCTCTGGAGG 840
QY 2021 GTGACATCCCTGAAGCTCGTGATGATCGTTGGCATGCGATGTTTACCATGACATGACA 2080
Db 841 GTGACATCCCTGAAGCTCGTGATGATCGTTGGCATGCGATGTTTACCATGACATGACA 900
QY 2081 GCTGGGGGAGGTCAATCGCAGGATTTGTTGCCAGCATCAATGAAGGATGACCCGCTGG 2140
Db 901 GCTGGGGGAGGTCAATCGCAGGATTTGTTGCCAGCATCAATGAAGGATGACCCGCTGG 960
QY 2141 TTCTCAGCTGCTATTTTCAGGATAGAGGACAGGAGCTGGTAGATGGGCTCAAAAGTCTGC 2200
Db 961 TTCTCAGCTGCTATTTTCAGGATAGAGGACAGGAGCTGGTAGATGGGCTCAAAAGTCTGC 1020
QY 2201 CTGAAGGGCTCTGAGGGCTTGAATAGCTGCAATGAGTACATGATGCCAGCCGATCATC 2260
Db 1021 CTGAAGGGCTCTGAGGGCTTGAATAGCTGCAATGAGTACATGATGCCAGCCGATCATC 1080
QY 2261 GTGTACCGGATGCTGAGAGACGGCAGCTGGAACACCTGCTGAGTACAGTACAGTACG 2320
Db 1081 GTGTACCGGATGCTGAGAGACGGCAGCTGGAACACCTGCTGAGTACAGTACAGTACG 1140
QY 2321 CAGTTTTTGGATTGTCTAAATCCATTGGTAGAGTTTACAACTTACAGTAACTT 2380
Db 1141 CAGTTTTTGGATTGTCTAAATCCATTGGTAGAGTTTACAACTTACAGTAACTT 1200
QY 2381 GTGGTGAAGAAAGATGAAACACCAAGATTTTGTCTCAGTCTGGAGAGAGATTCAGAAAT 2440
Db 1201 GTGGTGAAGAAAGATGAAACACCAAGATTTTGTCTCAGTCTGGAGAGAGATTCAGAAAT 1260
QY 2441 CCACCTTCTGGAAACAGTTATTGATGATGAGTTTACCAGACCAAGATGGTATGACTTTTTT 2500
Db 1261 CCACCTTCTGGAAACAGTTATTGATGATGAGTTTACCAGACCAAGATGGTATGACTTTTTT 1320
QY 2501 ATCGTGACCGAGCTGTGAGAGTGGTAGTGTCTCCACACATTAATGTCTAT 2560
Db 1321 ATCGTGACCGAGCTGTGAGAGTGGTAGTGTCTCCACACATTAATGTCTAT 1380
QY 2561 GACACAGCGGCTGAGCAGACCAATACAGCGCTTGACCTACAGCTGTGCCACATC 2620
Db 1381 GACACAGCGGCTGAGCAGACCAATACAGCGCTTGACCTACAGCTGTGCCACATC 1440
QY 2621 TATTACAACTGGCCAGGTGTCAATTCGTGTCTGCTGCTTGGCAGTACGCCACCAAGCTG 2680
Db 1441 TATTACAACTGGCCAGGTGTCAATTCGTGTCTGCTTGGCAGTACGCCACCAAGCTG 1500
QY 2681 GCTTTTCTGTTGGCCAGAGTATTCACAGAGGCCAAATCTGTCTCACTGTCAAAACCGCTT 2740
Db 1501 GCTTTTCTGTTGGCCAGAGTATTCACAGAGGCCAAATCTGTCTCACTGTCAAAACCGCTT 1560
QY 2741 TACTACTCTTAACCTCAGAGACGATGACGCGCTTTCTTTTGAATGACTTTGGGA 2800
Db 1561 TACTACTCTTAACCTCAGAGACGATGACGCGCTTTCTTTTGAATGACTTTGGGA 1620

QY 2801 TTTTAAAGCTTTTATTTTAACTTTTAACTGTTATCTTTCTGGATGAAACTTTGGGA 2860
Db 1621 TTTTAAAGCTTTTATTTTAACTTTTAACTGTTATCTTTCTGGATGAAACTTTGGGA 1680
QY 2861 AGGGGATTAAGGAGATCTAGCATTTTATTTCTAGCATTTCTTCAACCGCTTCCCTTATTT 2920
Db 1681 AGGGGATTAAGGAGATCTAGCATTTTATTTCTAGCATTTCTTCAACCGCTTCCCTTATTT 1740
QY 2921 TATATGTAATAAATTAAGATTTTATATTTTATCTTCTGTTCTCATAGATATTTTGTGAG 2980
Db 1741 TATATGTAATAAATTAAGATTTTATATTTTATCTTCTGTTCTCATAGATATTTTGTGAG 1800
QY 2981 CATTTTGTGTTTATTTTGAAGAAATGAGATCTTTGGTAGTATAAACAAGATCTC 3040
Db 1801 CATTTTGTGTTTATTTTGAAGAAATGAGATCTTTGGTAGTATAAACAAGATCTC 1860
QY 3041 TCTGAGAGTATTTGAAGATGCTTTGGAGATTTTAAACGTAATTTTCAAGAGTGAGCAA 3100
Db 1861 TCTGAGAGTATTTGAAGATGCTTTGGAGATTTTAAACGTAATTTTCAAGAGTGAGCAA 1920
QY 3101 GTCTTACTTATAAACCCTATATTAATTTTATTTTGGAGATCTTTTGAATTTTAAAGGA 3160
Db 1921 GTCTTACTTATAAACCCTATATTAATTTTATTTTGGAGATCTTTTGAATTTTAAAGGA 1980
QY 3161 GATAAGAGCGTAAAGTAGGATGCTACTACAACATAGTGGGTTTCAAGCTCATATCT 3220
Db 1981 GATAAGAGCGTAAAGTAGGATGCTACTACAACATAGTGGGTTTCAAGCTCATATCT 2040
QY 3221 TAAAGATAAAGGTAATTAATAAACCCTATACACAGATACAGAGAAATATGCTTGA 3280
Db 2041 TAAAGATAAAGGTAATTAATAAACCCTATACACAGATACAGAGAAATATGCTTGA 2100
QY 3281 TTTTATTTGGCAGGGGCTAGGTTGTATGGAGTAAAAAACAATTTTAA 3340
Db 2101 TTTTATTTGGCAGGGGCTAGGTTGTATGGAGTAAAAAACAATTTTAA 2160
QY 3341 ATTGTCCAAAGAAACATTTTAAAGTCTTTTAAACAAAGCCATGATTAATCTCTATA 3400
Db 2161 ATTGTCCAAAGAAACATTTTAAAGTCTTTTAAACAAAGCCATGATTAATCTCTATA 2220
QY 3401 TTAACATCAGCTATTTTATTTTGGAACTGGGACATGATCTTATTTGTATATAAATAA 3460
Db 2221 TTAACATCAGCTATTTTATTTTGGAACTGGGACATGATCTTATTTGTATATAAATAA 2280
QY 3461 AATTGATGTCCC 3472
Db 2281 AATTGATGTCCC 2292

RESULT 3

US-09-873-737A-3
; Sequence 3, Application US/09873737A
; Patent No. US20020076797A1
; GENERAL INFORMATION:
; APPLICANT: Duke University
; APPLICANT: Lin, Haifan
; TITLE OF INVENTION: PURIFIED AND ISOLATED piwi FAMILY GENES AND GENE
; FILE OF INVENTION: PRODUCTS AND THERAPEUTIC AND SCREENING METHODS USING SAME
; FILE REFERENCE: Attorney Docket No. US20020076797A1 180-104/2
; CURRENT APPLICATION NUMBER: US/09/873, 737A
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US99/28764
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/110, 901
; PRIOR FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 4064
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: CDS

Query Match 15.9%; Score 550.4; DB 10; Length 590;
Best Local Similarity 99.8%; Pred. No. 7.2e-115;
Matches 551; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2915 TTATTTATATGTAATAATTAAGATTTTATATTTTATCTTCTGTTCTCATAGATTTT 2974
DB 590 TTATTTATATGTAATAATTAAGATTTTATATTTTATCTTCTGTTCTCATAGATTTT 531
QY 2975 TGTGAGCAATTTTGTGTTTATTTTGAAGAAATGCGATTAAGATCTGTAGATTAAC 3034
DB 530 TGTGAGCAATTTTGTGTTTATTTTGAAGAAATGCGATTAAGATCTGTAGATTAAC 471
QY 3035 AGACTCTCTAGAGATTTTGAAGATGCTTTGAGATTTACTTAAACGATCTTTCAGAGT 3094
DB 470 AGACTCTCTAGAGATTTTGAAGATGCTTTGAGATTTACTTAAACGATCTTTCAGAGT 411
QY 3095 GAGCAAGTCTTACTTAAACCTATATTAATTTTATTTTGAAGATCTGTTTGAATTT 3154
DB 410 GAGCAAGTCTTACTTAAACCTATATTAATTTTATTTTGAAGATCTGTTTGAATTT 351
QY 3155 AAAGAGATTAAGAGCGGTAAAGTAGATGCTCACTAACCATAGGTGGGTTTCAGCTC 3214
DB 350 AAAGAGATTAAGAGCGGTAAAGTAGATGCTCACTAACCATAGGTGGGTTTCAGCTC 291
QY 3215 ATATCTTAAAGATTAAGAGCTATATTAATTAATTAACCTATACAGAGATTAAGAGAAATAT 3274
DB 290 ATATCTTAAAGATTAAGAGCTATATTAATTAATTAACCTATACAGAGATTAAGAGAAATAT 231
QY 3275 GCTTGAATTTTATTTTGAAGAGGCGGTAGGTTGATGAGGATTAAGAGAAATAT 3334
DB 230 GCTTGAATTTTATTTTGAAGAGGCGGTAGGTTGATGAGGATTAAGAGAAATAT 171
QY 3335 TTTTAAATTTGTCAAAGAAACATTTTAAAGCTTTTAAACAAAAGGCGCATGATTAATC 3394
DB 170 TTTTAAATTTGTCAAAGAAACATTTTAAAGCTTTTAAACAAAAGGCGCATGATTAATC 111
QY 3395 TCTATATTAACATCACTATTTATTTTGTGAACTGGGACATGATTTATTTGTATTA 3454
DB 110 TCTATATTAACATCACTATTTATTTTGTGAACTGGGACATGATTTATTTGTATTA 51
QY 3455 AAATAAATTTGATGT 3469
DB 50 AAATAAATTTGATGT 36

RESULT 6
US-09-920-300A-34
Sequence 34, Application US/09920300A
Patent No. US20020136728A1
GENERAL INFORMATION:
APPLICANT: King, Gordon E.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Xu, Jiangchun
APPLICANT: Secretist, Heather
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.547
CURRENT APPLICATION NUMBER: US/09/920.300A
NUMBER OF SEQ ID NOS: 1789
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 34
LENGTH: 501
TYPE: DNA
ORGANISM: Homo sapiens
US-09-920-300A-34

Query Match 13.2%; Score 459.4; DB 10; Length 501;
Best Local Similarity 99.8%; Pred. No. 2.3e-94;
Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2695 CCAGAGTATTCACAGAGCAAAATCTGTCACTGCAACCGCTTTACTACTTAAAC 2754
DB 460 CCAGAGTATTCACAGAGCAAAATCTGTCACTGCAACCGCTTTACTACTTAAAC 2754

DB 1 CCAGAGTATTCACAGAGCAAAATCTGTCACTGCAACCGCTTTACTACTTAAAC 60
QY 2755 TGCAGAAAGCATGAGAGCGGCTTTTCTTTTGAATGACCTTGGATTTTAAAGCTTT 2814
DB 61 TGCAGAAAGCATGAGAGCGGCTTTTCTTTTGAATGACCTTGGATTTTAAAGCTTT 120
QY 2815 TATTTACTTTTCTTTTAACTGTTATCTTTCGATGAAACCTTGGAGGAGATTAAGAGA 2874
DB 121 TATTTACTTTTCTTTTAACTGTTATCTTTCGATGAAACCTTGGAGGAGATTAAGAGA 180
QY 2875 TCTACATTTTATTTTCTGACATTTGCTATTCACCGGCTTCTTATTTATATGTAATAAT 2934
DB 181 TCTACATTTTATTTTCTGACATTTGCTATTCACCGGCTTCTTATTTATATGTAATAAT 240
QY 2935 AAGATTTTATTTTATTTCTTCTGTTTCTCAATGATTTTGGAGCAATTTTGTGTTA 2994
DB 241 AAGATTTTATTTTATTTCTTCTGTTTCTCAATGATTTTGGAGCAATTTTGTGTTA 300
QY 2995 TTTTGAAGAAATGAGATTAAGATCTGTAGTAAACAGACTCTCTGAGATTTTG 3054
DB 301 TTTTGAAGAAATGAGATTAAGATCTGTAGTAAACAGACTCTCTGAGATTTTG 360
QY 3055 AAATGCTTTGAGATTTTACTTAAACGATCTTTCAGAGTGAAGCAAGTCTTACTTAA 3114
DB 361 AAATGCTTTGAGATTTTACTTAAACGATCTTTCAGAGTGAAGCAAGTCTTACTTAA 420
QY 3115 CCTATATTAATTTATTTTGAAGTACCTGTTTGAATTTA 3155
DB 421 CCTATATTAATTTATTTTGAAGTACCTGTTTGAATTTA 461

RESULT 7
US-10-033-528-34
Sequence 34, Application US/10033528
Patent No. US20020131971A1
GENERAL INFORMATION:
APPLICANT: King, Gordon E.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Xu, Jiangchun
APPLICANT: Secretist, Heather
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.547C1
CURRENT APPLICATION NUMBER: US/10/033.528
NUMBER OF SEQ ID NOS: 1896
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 34
LENGTH: 501
TYPE: DNA
ORGANISM: Homo sapiens
US-10-033-528-34

Query Match 13.2%; Score 459.4; DB 12; Length 501;
Best Local Similarity 99.8%; Pred. No. 2.3e-94;
Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2695 CCAGAGTATTCACAGAGCAAAATCTGTCACTGCAACCGCTTTACTACTTAAAC 2754
DB 1 CCAGAGTATTCACAGAGCAAAATCTGTCACTGCAACCGCTTTACTACTTAAAC 60
QY 2755 TGCAGAAAGCATGAGAGCGGCTTTTCTTTTGAATGACCTTGGATTTTAAAGCTTT 2814
DB 61 TGCAGAAAGCATGAGAGCGGCTTTTCTTTTGAATGACCTTGGATTTTAAAGCTTT 120
QY 2815 TATTTACTTTTCTTTTAACTGTTATCTTTCGATGAAACCTTGGAGGAGATTAAGAGA 2874
DB 121 TATTTACTTTTCTTTTAACTGTTATCTTTCGATGAAACCTTGGAGGAGATTAAGAGA 180
QY 2875 TCTAGCAATTTATTTCTGACATTTGCTATTCACCGGCTTCTTATTTATATGTAATAAT 2934
DB 181 TCTAGCAATTTATTTCTGACATTTGCTATTCACCGGCTTCTTATTTATATGTAATAAT 240

QY 2935 AAGATTTTATATTTATCTTCTGTTTCTCATAGATATTTTGTGAGCATTTTGTGTTT 2994
 DB 241 AAGATTTTATATTTATCTTCTGTTTCTCATAGATATTTTGTGAGCATTTTGTGTTT 300
 QY 2995 TTTTGAAGAAATGTTGATAGATATCTGTTAGTATATAAACAAGACTCTCTGAGATTTTG 3054
 DB 301 TTTTGAAGAAATGTTGATAGATATCTGTTAGTATATAAACAAGACTCTCTGAGATTTTG 360
 QY 3055 AAATGTTTGGAGATTTTAACTTAAAGTACTTTCAGGAGTGAAGCTCTTACTTATATAA 3114
 DB 361 AAATGTTTGGAGATTTTAACTTAAAGTACTTTCAGGAGTGAAGCTCTTACTTATATAA 420
 QY 3115 CCTATATTAACCTTTATTTTGGAGATACCTGTTTGAATTTA 3155
 DB 421 CCTATATTAACCTTTATTTTGGAGATACCTGTTTGAATTTA 461
 RESULT 8
 US-09-777-564-1706/c
 ; Sequence 1706, Application US/09777564
 ; Patent No. US20020022591A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Mannion, Jane
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
 ; FILE REFERENCE: 210121.493
 ; CURRENT APPLICATION NUMBER: US/09/777,564
 ; PRIOR FILING DATE: 2001-02-05
 ; NUMBER OF SEQ ID NOS: 1730
 ; SOFTWARE: FastSeq for Window Version 4.0
 ; SEQ ID NO 1706
 ; LENGTH: 460
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-777-564-1706

Query Match 13.2%; Score 458.4; DB 10; Length 460;
 Best Local Similarity 99.8%; Pred. No. 3.7e-94;
 Matches 459; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2695 CCAGAGTATTCACAGAGAGCAATCTGCTACCTGCTCAACCGCTTTTACTACCTCTAAC 2754
 DB 460 CCAGAGTATTCACAGAGAGCAATCTGCTACCTGCTCAACCGCTTTTACTACCTCTAAC 401
 QY 2755 TGCAAGAGACGATCGACCGCTTTTCTTTTGAATGACCTTTGGGATTTTAAAGCTTT 2814
 DB 400 TGCAAGAGACGATCGACCGCTTTTCTTTTGAATGACCTTTGGGATTTTAAAGCTTT 341
 QY 2815 TATTTACTTTTTTTTAACTGTTATCTTCTGATGAACTTGGGATTTTAAAGCTTT 2874
 DB 340 TATTTACTTTTTTTTAACTGTTATCTTCTGATGAACTTGGGATTTTAAAGCTTT 291
 QY 2875 TCTAGCATTTTATTTCTAGCATTTCTATTCACCGCTTCTTATTTATATGTAATAAAT 2934
 DB 280 TCTAGCATTTTATTTCTAGCATTTCTATTCACCGCTTCTTATTTATATGTAATAAAT 221
 QY 2935 AAGATTTTATATTTATCTTCTGTTTCTCATAGATATTTTGTGAGATTTTGTGTTT 2994
 DB 220 AAGATTTTATATTTATCTTCTGTTTCTCATAGATATTTTGTGAGATTTTGTGTTT 161
 QY 2995 TTTTGAAGAAATGTTGATAGATATCTGTTAGTATATAAACAAGACTCTCTGAGATTTTG 3054
 DB 160 TTTTGAAGAAATGTTGATAGATATCTGTTAGTATATAAACAAGACTCTCTGAGATTTTG 101
 QY 3055 AAATGTTTGGAGATTTTAACTTAAAGTACTTTCAGGAGTGAAGCTCTTACTTATATAA 3114
 DB 100 AAATGTTTGGAGATTTTAACTTAAAGTACTTTCAGGAGTGAAGCTCTTACTTATATAA 41
 QY 3115 CCTATATTAACCTTTATTTTGGAGATACCTGTTTGAATTT 3154
 DB 40 CCTATATTAACCTTTATTTTGGAGATACCTGTTTGAATTT 1

RESULT 9
 US-09-801-574-15
 ; Sequence 15, Application US/09801574
 ; Patent No. US20020081592A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Peijing Jeremy
 ; APPLICANT: Page, David C.
 ; TITLE OF INVENTION: Reproduction-Specific Genes
 ; FILE REFERENCE: 0399.2007-002
 ; CURRENT APPLICATION NUMBER: US/09/801,574
 ; CURRENT FILING DATE: 2001-03-07
 ; PRIOR APPLICATION NUMBER: 60/187,518
 ; PRIOR FILING DATE: 2000-03-07
 ; PRIOR APPLICATION NUMBER: 60/261,557
 ; PRIOR FILING DATE: 2001-01-12
 ; NUMBER OF SEQ ID NOS: 90
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 15
 ; LENGTH: 3649
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; US-09-801-574-15

Query Match 10.9%; Score 377; DB 10; Length 3649;
 Best Local Similarity 52.5%; Pred. No. 2.2e-75;
 Matches 974; Conservative 0; Mismatches 862; Indels 21; Gaps 6;
 QY 914 CCTGGCTTCACTACTTCCATCTTTCAGTATGAAACAAGCATCATGCTCTGCACTGAGTT 973
 DB 1 CCTGGCTATGCGGTAGTATCCGAGGACAGACGGGGTCTCTTCTGCTCGTGAATGC 60
 QY 974 AGCCATTAAGTCTTTCGAAGTGAGACTGTTTGGATTTTCACTTTTATCATCAG 1033
 DB 61 TCTATAAGTTCATTCGGAACGACTCTGTGCTGATGCTCATGATGCTATCTACCCAG 120
 QY 1034 ACAGAAGAACATAAATTTCAAGAACAAAGTTTCCAAAGAAATAATAGTTTGTGTTCT 1093
 DB 121 AACAGGAGCAC---TTCCAGGACGAGTGCGAGCAAGCTTCTGGTTGGCAGCATTTGTCATC 177
 QY 1094 ACCAAGTATAACATTAAGACATACAGAGTGGATGATATTTCACTGGGACAGAAATCCCAAG 1153
 DB 178 ACGGCTACAAACAATCGTACCTACCGAATCGATGATGAGCTGGAAACAAGACCCCTTAA 237
 QY 1154 AGCACTTTTAAAGAAAGCGACGCTCTGGGCTCAGCTTCTTGAATATCTACAGGAAGCAA 1213
 DB 238 GACAGCTTTGTCATGTCGAGCGGGAAGAAATACATTTCTTGAATATCTACAGCAAAAC 297
 QY 1214 TACAACCAAGAGATCACCGACTTGAAGCAGCTCTGTTGGTCAAGCCAGCCCA---AGAGA 1270
 DB 298 TATGGGATCACAGTCAAGGAAGATGACCAAGCGCTGCTGATCCACCGCCAGTGAGAGA 357
 QY 1271 AGCGGGGCTTGGGGGACACTGCCAGGCTGCCATGCTCATTTCTGAGCTCTGCTAT 1330
 DB 358 CAGATAACCATGGGATGTTGCTGAAGGCGAGATCTCTGCTGCTGCGGAGCTCTCTTC 417
 QY 1331 CTTACAGCTCTAACTGATAAATAATGCGTAATGATTTTAACTGATGAAAGACTTACCGCTT 1390
 DB 418 ATGACGGGATCCCTGAGAGATGAAGAGGACTTTCAGGCGCATGAAGGACTTCACTCAG 477
 QY 1391 CATACAAGACTAACTCCAGAGCAAGGCGTGAAGTGGGACGACTCATTTGATTACATT 1450
 DB 478 CAGATTAACCTTGAGCCCAAGCAGCACCAAGCGTGTCTTGGAAATGCTCTCTGAGAGATT 537
 QY 1451 CATAAAAACGATAATGTTTCAAGGGAGCTTCAGAGCTGGGGTTTGGATTTTGTATCCAAAC 1510
 DB 538 TCACAAAACGACAGACGAGCAATGAGCTGCCGCTGGGGCTCAGTCTGCTGATAAAGAT 597
 QY 1511 TTACTGCTCTTCTCAGGAAGAAATTTTGAACAAGAAAGATTCACCAAGGTGGAAAAACA 1570
 DB 598 GTCCACAAGATGAAGGTGCGGCTTCTGCAATGGAGAGGATCACTTAAGGAACAACCTTCA 657
 QY 1571 TTTGATTACATCCACAAATTTTCAGATTTGGTCCAAAGAAAAAAGAGGTGCACCAATAATT 1630

Db 658 TTGTG---CACATCGAGGCGCTGAACTGGGTTAAGAGTGAACCAAGATGCTTCATT 714
 Oy 1631 AGGTGTAAGCCCTGATGTAATGCTGTTGATCTATACGCAAGAAATTATGAGCAGCC 1690
 Db 715 CTACTATTTCCATGATCTTTGCGACCTTTTATCCAAAGAGCAATGACCAAGCC 774
 Oy 1691 AATTGATGATGTAATGTAATTTAATTTAATGTAACACAGCCTAGGCAATGCAAAATGAGAAA 1750
 Db 775 AGAGATCGGTTAATGATTTGAGAAAGATTTGCGGCGCCATGCGATGCGCAAGCCCC 834
 Oy 1751 GCATATATGTTAAAGT---GGATGACAGACCTGAAGCTTAATGAGTCTTACAGCA 1807
 Db 835 CCGAGCTGGGTGAGCTGAAGATGACCGAATGAGACCTTATATGAGACCATTCAGTCC 894
 Oy 1808 AAGGTACAGCAGAC-----ACCGAATGATGTTGCTGCTGTTGTCAGATGTAATGCGAAG 1861
 Db 895 TTACTGGAGTTGAGGAGAGATACAAATGCTGTTGATCATCATGAGGCGACAGTGAAT 954
 Oy 1862 GACAATATGATGCTATTAATAAATACCTGTTGATGAGATGCTTACCCCAAGTGT 1921
 Db 955 GATCTATGAGAGCATCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014
 Oy 1922 GTGTGCGCCGAACTTGAAGCAACAGCAAACTGTCATGAGCCATTGCTACAAAGTTGCC 1981
 Db 1015 ATCATATGCGAATCATTTGGTCAAGCCACAGGCTTGGAGCTGAGTGGCTGAGAAATTTTA 1074
 Oy 1982 CTACAGATGAACTGCAAGATGAGGAGAGAGCTGAGAGGAGTGAATCCCTGAGACTC 2041
 Db 1075 CTTCAGATGAACTGTAACCTGAGGAGTGAAGCTGAGGAGTGAATTTCCCTGAGAAACA 1134
 Oy 2042 GTGATGATGCTGAGATGATGTTGTTACATGATGATGATGATGATGATGATGATGATGAT 2101
 Db 1135 CTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1194
 Oy 2102 GGATTTGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2161
 Db 1195 GGTCTGAGGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1254
 Oy 2162 GATGAGGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2221
 Db 1255 ATGCGCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1314
 Oy 2222 TGGAT 2281
 Db 1315 TATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1374
 Oy 2282 GACGCGCAGCTGAAACACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2341
 Db 1375 GATGCGCAGCTGAAACACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1434
 Oy 2342 TCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2401
 Db 1435 GCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1491
 Oy 2402 ACCAGATTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2461
 Db 1492 ACCAGATTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1551
 Oy 2462 GAT 2521
 Db 1552 GAT 1611
 Oy 2522 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2581
 Db 1612 CAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1671
 Oy 2582 GACCAATACAGCCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 2641
 Db 1672 GAT 1731
 Oy 2642 ATTGCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2701

Db 1732 ATCCGAGTTCCAGCTCTCTGAGATGATGCCACAGAGTACTTTCTGTCGAGACAGATT 1791
 Oy 2702 ATTCACAGAGCCAAATCTGTACTGTCAACCGCTTTACTACCTTAACCTGCA 2758
 Db 1792 TTGCTATGATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1848

RESULT 10

US-09-873-737A-1
 ; Sequence 1, Application US/09873737A
 ; Patent No. US20020076797A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Duke University
 ; APPLICANT: Liu, Haifan
 ; TITLE OF INVENTION: PURIFIED AND ISOLATED P1W1 FAMILY GENES AND GENE
 ; TITLE OF INVENTION: PRODUCTS AND THERAPEUTIC AND SCREENING METHODS USING SAME
 ; FILE REFERENCE: Attorney Docket No. US20020076797A1 180-104/2
 ; CURRENT APPLICATION NUMBER: US/09/873,737A
 ; PRIOR FILING DATE: 2001-06-04
 ; PRIOR APPLICATION NUMBER: PCT/US99/28764
 ; PRIOR FILING DATE: 1999-12-03
 ; PRIOR APPLICATION NUMBER: 60/110,901
 ; NUMBER OF SEQ ID NOS: 21
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 3047
 ; TYPE: DNA
 ; ORGANISM: Drosophila sp.
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (84)..(2612)
 ; NAME/KEY: misc feature
 ; LOCATION: (120)
 ; OTHER INFORMATION: n=a or c, Xaa=Leu or Ile
 ; NAME/KEY: misc feature
 ; LOCATION: (399)
 ; OTHER INFORMATION: n=a or t, Xaa=Leu or Ile
 ; NAME/KEY: misc feature
 ; LOCATION: (2436)
 ; OTHER INFORMATION: n=a or c, Xaa=Leu or Ile
 ; US-09-873-737A-1

Query Match 7.3%; Score 253; DB 10; Length 3047;
 Best Local Similarity 47.1%; Pred. No. 1.9e-47;
 Matches 1080; Conservative 0; Mismatches 1184; Indels 29; Gaps 9;

Oy 470 GATCAAAAAGAGTTCTTCAAGCATTTATGTAAGTTAAGCTTAACCATTTCCGGCTG 529
 Db 342 GATTCAGAGAGAGGAAACGATGGGCTCCGGCTCATGCTGCAGAGAACTTTTTCAGANTA 401
 Oy 530 AATCCCGTCCCGAGTGGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 589
 Db 402 AAAACCAAGCCGAAATGCGAGTCTTCAATTAATGATGATGATGATGATGATGATGATGATGAT 461
 Oy 590 GCAGAAAGATCCGTTAGCTCTTTTCAACGAGAACTTAATGTAAGTTGAAAGTTCAT 649
 Db 462 AATCTCTGCTGCTGATGAGATTTTTCATATGATGATGATGATGATGATGATGATGATGATGAT 521
 Oy 650 GCTTTGATGAGATGATTTATTTTACCTAAAGACTACAGCAAAAGTTTGAAGTT 709
 Db 522 CTATTCGAGAGAGTGCATGCTTTCACACAGCAAAATGAGAGAAATCAGCGTGTCTC 581
 Oy 710 TTATGTAAGACCCGAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 769
 Db 582 AGCGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 641
 Oy 770 CACATCAATCAACAACTTTTGAAGTCTTAATATTTATTTTGAAGGCTTTTGAA 829
 Db 642 TCGTGTGCTGAGCCGCTTTTGAAGTCTTAATATTTTGAAGGCTTTTGAAAG 701
 Oy 830 ATCATGATTTGCAAAATTTGAGCAGAAATTTATTAACCAATGATGCCAATTTGATTT 889

Db 702 GGCCTAAATTTGGAAATTTAGTTGGCCGTAATCTCTTTGATCCCGAGCTAAGATCGAAATA 761
 Qy 890 CCAAGTCAAGTTGGTGAATTTGGCTGGCTTCACTCTTCATCTCTTCAGTATGAAAC 949
 Db 762 AGGAGTTCAAAATGGAGCTATGGCCGGCTTATGAGACATCGATTCGTGACGAGAA 821
 Qy 950 AGCATCATGCTCTGCACTGAGCTTGGCCATGAAGTCTTCCGAGTGAAGTCTTTGGAT 1009
 Db 822 GATAATTTATTTGGGCAACGAATTAATCTCAAAAGTTATGGCCAGGAGCATCTACGAC 881
 Qy 1010 TTGATGTTCAACTTTTATCATGACAGAGAAACATAAATTTTCAAGAAACAAGTTTCCAAA 1069
 Db 882 ATAATGCGAGCTTGGCTCACCAATCCGGCTCGTCATCAGGACGAAGTACGGGT- - -A 935
 Qy 1070 GAATAATAGCTTTAGTTGTTCTTACCAAGTATAACATTTAAGACATACAGATGATGAT 1129
 Db 936 AATGTTTGGACTTGAATTTGCTTACCGATTTACGATTTAACAATACAGAACTTTATCGTATCAATGAT 995
 Qy 1130 ATTGACTGGGACAGAAATCCCAAGAGACCTTTTAAAGAAAGCGAGCTCTGGGTGAGC 1189
 Db 996 GTGACTTTGGACAAACTCCGAAATCAACATTT- - -CAGTTGCAAGGTAGAGATACGT 1052
 Qy 1190 TTCTTGAATACACTACAGGAGCAATCAACCAAGAGATCACCGACTTGAAGCAGCTGTC 1249
 Db 1053 TTGCTGATGATCTCTCACTAAATATAATATACGCAATTCGCGACCAACAATCAGCGCTG 1112
 Qy 1250 TTGCTGAGCCAGCCCAAGAGAGGGGGCCCTGGGGGAGACTCCAGGGCCCTGCCATG 1309
 Db 1113 CTGATTTCCAAA- - -AATAGGACAGGCTCTAATAACTAAGCTTACGCTAGGAAATAGTGTA 1169
 Qy 1310 CTGATCTGAGCTCTGCTATCTTACAGTCTAATGATTAATAAATGCGTAAATTTTAAAC 1369
 Db 1170 CTAATTTCTGAGCTCTGCGAGTGTGCTGAGTCAATGCGAGATGCGCTCAAACTTTTAC 1229
 Qy 1370 GTGATGAAGAATTAGCCGCTTATACAGACTTACAGTCTAATGATTAATAAATGCGTAAATTTTAAAC 1429
 Db 1230 CTTATGCGTGCATGAGCAGTATACGCAATGAGCAATGAGCAATGAGCAATGAGCAATG 1289
 Qy 1430 GGAAGACTCATTTGATTAATCATTAATAAAGATTAATGTTTCAAGGGAGCTTCCAGACTGG 1489
 Db 1290 CGCGCTTTTAAACACCGTTTCAAAACACTCCAGAAAGTGTGAGGCTTTGAGAGACTGG 1349
 Qy 1490 GGTGTTGAGCTTTGATTTCAACTTACTGCTTCTCAGGAGAAATTTTGCAAAAGAGAG 1549
 Db 1350 AACATGGAACCTGGAAGAGAGCTCAGAGAGTACAGGCGCGGATTAATTTGAGACAGAAAC 1409
 Qy 1550 AT- - -TCACCAAGTGGAAAAACAATTTGATTAACAATCCACAATTTGAGATTTGGTCCAA 1605
 Db 1410 ATCGTGTTCATAATGGAAGGTTCTCTGCTGGAGAAACCGCTGATTTGGCAAGGCACTTC 1469
 Qy 1606 AGAAACAGAGGTGCACCAATTAATGATTTAAGCCACTAGATAAATGCTGCTGTTGATCTA 1665
 Db 1470 AGAGCCAAAGAGTCTTACCACTCCAGAGGATGGC-CTCGATCGTTGGGCTGTATCGC 1528
 Qy 1666 TACGGAAGAAATTTAAGACGAGCAATTCATTTGATACAAAATCTATTTAAAGTTTACACC 1725
 Db 1529 GCCGCAAGGAAATTCCTCAACTCCGAATCTTACTTGTACTCTTTGATAGAGAGCTAG 1588
 Qy 1726 AGCCATGGCATGCAATGAGAAAGC- - -ATAATGATTTGAAGTGGATGACAGAACTGA 1782
 Db 1589 TGGATGGGTCTTAGAATTCGAAGCCCGCCAGGAAATTCATAATTTATGATGATCGCACTGG 1648
 Qy 1783 AGCCTACTTAAGAGTCTTACAGCAAAAGGTTCACAGCAGACCCAGATAGTTGCTGCT 1842
 Db 1649 AACTTATGTGAGAGCAATGGATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1708
 Qy 1843 GTTGTCAAGTAAATCGGAAGCAAAATACGATGCTTATTAATAAATACCTGTGTACAGATTG 1902
 Db 1709 CGTACCAATGATTAACCGCGGAAAGATCTCATCAATCAAAAGAGAGGATACGTTGACAG 1768
 Qy 1903 CCCTACCCCAAGTCACTGTGTGTCGCGGACCTTAGGCAACAGCAAACTGTCTATGGC 1962
 Db 1769 GCGGTGCCAACTCA- - -AGTTGTGACCTTTAAACGACCAAGAACCGTAGCCTTATGAG 1825

Qy 1963 CATTGCTACAAAGATTGCGCCTACAGATGAATGCAAGATGGGAGGAGAGCTCTGCGAGGT 2022
 Db 1826 CATTGCGCACCAAAATAGCAATCCAACTGAATTTGCAAGTTGGGATATACACCTTGATGAT 1885
 Qy 2023 GGACATCCCCCTCAAGCTCGTGTATGATCGTTGGCATCGATTTGTTACCATGATGACAGC 2082
 Db 1886 CGAATCACTCTGTCGGACTGATGACAAATTTGGCTTTGACATTTGCAAGAGACACAGAGA 1945
 Qy 2083 TGGCGGAGGTCAATCGCAGGATTTGTCGAGCATCAATGAAGGATGACCCGCTGTT 2142
 Db 1946 TCGGAAGAGGGCTTACGAGCATTTGATTTGCTCAATGATCTACAGCAAACTCCAGCTA 2005
 Qy 2143 CTCACCTCGATATTTCAAGGATAGAGACAGAGCTGGTGTAGATGGGTCTAAAGTCTGCT 2202
 Db 2006 CTTGACACAGTCAAGGAGTGGAGCGCTTTGATGCTGCTAAACACCTTTGCGCGAT 2065
 Qy 2203 GCAAGC- - -GGCTCTGAGGGCTTTGGAATAGCTGCAATGAGTACATGCCAGCGGATCAT 2259
 Db 2066 GATAGCAAGGCGCTTGGCAATATCAACATGACATAGGAAGCTGCTCGATCTCGAATCGT 2125
 Qy 2260 CGTGTACCGCATGCGGTAGGAGCGGCGAGCTGAAACACTGGTGAATACGAAAGTGC 2319
 Db 2126 ATTTTATCGAGCGGTGAGCTCGGCTCTCTAAAGCAGCTTTTGAATTTGAAGTCAA 2185
 Qy 2320 ACAGTTTTTGGATTTCTTAAATCCATTTGGTAGAGTTTACAACTAGANTAAACGGTAAT 2379
 Db 2186 GGACATCATTTGAGAAGTTGAAACTGAACTAGCCCGCTCCAGCTAAGCCCAACCGCAAT 2245
 Qy 2380 TGTGTTGAAGAAAGAGTGAACACAGATTTTGTCTGCTGAGGAGAGACTTCAGAA 2439
 Db 2246 AGCT- - -TATATTTGTTGTAACAGATCCATGAACACCGCTTCTTCTCAACGACAAA 2302
 Qy 2440 TCCACTTCTCGGAACAGTTATGATGTAGAGTTTACAGACCAAGATGATGATGATGAT 2499
 Db 2303 TCTCTCGCTGTTACTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2362
 Qy 2500 TATGTTGAGCGAGCTGTGAGAAAGTGTGATGTTTCTCCACACATTTACAATGATCTA 2559
 Db 2363 TCTGCTCTCGCAACAGTTCTGTCAGGCTGAGTGTGCGGACGAGCTACATGTTCTTTA 2422
 Qy 2560 TGACAAACAGCGGCTGAGCCAGACACATACAGCGCTTGAACCTCAAGCTGTGCCAT 2619
 Db 2423 TAGCAGATCGGTTCTTCCCGGAGAAATGCAAAACTTACGTACAAGATGTGCCACT 2482
 Qy 2620 CTATTAACACTGCCAGGTGTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2679
 Db 2483 GTACTACAATTTGTCGGGACACACAGCTGCGAGGTTTGGCAGTACGCTTAAGAAGCT 2542
 Qy 2680 GGCTTTTCTGTTGGCCAGAGTATTCACAGAGGCAAAATCTGTCTGTCAAAACCGCT 2739
 Db 2543 AGTACCTCTGTTGGTACAACTTGTGACTTATTCGCAAAACCGCTCGAAAGAGTT 2602
 Qy 2740 TTACTACCTTAA 2752
 Db 2603 TTATTATCTATTA 2615

RESULT 11
 US-09-774-434-2
 ; Sequence 2, Application US/09774434
 ; Patent No. US20020137044A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Goodrich, Ryle
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Drmanac, Radoje T.
 ; TITLE OF INVENTION: No. US20020137044A1el Nucleic Acids and
 ; TITLE OF INVENTION: Polypeptides

```

FILE REFERENCE: 787CIP2G
CURRENT APPLICATION NUMBER: US/09/774,434
CURRENT FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 6
SOFTWARE: pf_genes Version 2.0
SEQ ID NO 2
LENGTH: 2914
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(2583)
US-09-774-434-2

```

```

Query Match 3.2%; Score 111.8; DB 10; Length 2914;
Best Local Similarity 50.5%; Pred. No. 1.3e-15;
Matches 306; Conservative 0; Mismatches 288; Indels 12; Gaps 1;

```

```

QY 2093 TCAATCGAGATTGTTGTCAGATCAATGAGGATGACCGCGTGTCTCAGCGCTGC 2152
DB 1831 TCTATTGCTGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1890
QY 2153 ATATTTGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2212
DB 1891 AGAGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1950
QY 2213 CTGAGGCGCTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2272
DB 1951 CTATTTCAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2010
QY 2273 GGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2332
DB 2011 GGTGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2070
QY 2333 TGTCTAAATTCATTGTTAGAGGTTACACCTTGAATTAAGGTAATTTGTTGGAAGAA 2392
DB 2071 GCGTGATGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2130
QY 2393 AGAGTGAACACGAGATTTTGTCTGATGATGATGATGATGATGATGATGATGATG 2440
DB 2131 AGACATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2190
QY 2441 CCACTTCTGGAACAGTTATGATGATGATGATGATGATGATGATGATGATGATG 2500
DB 2191 ATCCAGCTGGAACAGTTATGATGATGATGATGATGATGATGATGATGATGATG 2250
QY 2501 ATGTGAGCAGGCTGTGAGAGTGTAGTGTGTCTCCACACATTAATGATCTAT 2560
DB 2251 CTCTGAGCAGCTGTGAGAGTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2310
QY 2561 GACAAAGCGGCTGGAAGCAGACATACAGCGGTTGATGATGATGATGATGATG 2620
DB 2311 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2370
QY 2621 TATTACATGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATG 2680
DB 2371 TACGTACGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2430
QY 2681 GCTTTT 2686
DB 2431 GCATTT 2436

```

```

RESULT 12
US-09-815-343-3
Sequence 3, Application US/09815343
Patent No. US20010055596A1
GENERAL INFORMATION:
APPLICANT: Meagher, Madeleine

```

```

APPLICANT: Xu, Jiangchun
APPLICANT: King, Gordon E.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
FILE REFERENCE: 210121.504
CURRENT APPLICATION NUMBER: US/09/815,343
CURRENT FILING DATE: 2001-03-22
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 444
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(444)
OTHER INFORMATION: n = A,T,C or G
US-09-815-343-3

```

```

Query Match 2.8%; Score 95.6; DB 10; Length 444;
Best Local Similarity 66.2%; Pred. No. 2.5e-12;
Matches 137; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

```

```

QY 1824 CCAGATAGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1883
DB 122 CCAGATAGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 181
QY 1884 AATCTGTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1943
DB 182 AATTTTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 241
QY 1944 AACGCAAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2003
DB 242 AACGCGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 301
QY 2004 GAGGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2030
DB 302 GAGGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 328

```

```

RESULT 13
US-09-864-761-32092
Sequence 32092, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665

```

```

? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00668
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00663
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00662
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00661
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00670
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: US 60/234,687
? PRIOR FILING DATE: 2000-09-21
? PRIOR APPLICATION NUMBER: US 09/508,408
? PRIOR FILING DATE: 2000-06-30
? PRIOR APPLICATION NUMBER: US 09/774,203
? PRIOR FILING DATE: 2001-01-29
? NUMBER OF SEQ ID NOS: 49117
? SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
? SEQ ID NO 32092
? LENGTH: 213
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: MAP TO AP000359.1
? OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
? OTHER INFORMATION: EST HUMAN HIT: BF246708.1, EVALUATION 5.00e-05
? OTHER INFORMATION: NT HIT: G11094611, EVALUATION 8.00e-06
? OTHER INFORMATION: SWISSPROT HIT: O04379, EVALUATION 2.70e-02
? US-09-864-761-32092

```

Query Match	2.3%	Score	78.6;	DB	10;	Length	213;
Best Local Similarity	67.2%	Pred.	No. 1.2e-08;				
Matches 127;	Conservative	0;	Mismatches	59;	Indels	3;	Gaps
QY	1832	GTGTCTGTCCTCTCGCAAGTAATCGGAGGACAAATACGATGCTATTAAAAATACCTTG	1891				
Db	28	GTGATTTGTATCTCGCCCAATGATGACAAACGTAAGATATGACAGCATAAAGAATACCTA	87				
QY	1892	TGTA CAGATTGCCCTACCCCAAGTCAGTGTGTGTGGTCCCGCAACCTTACGCCAAACAGCAA	1951				
Db	88	TGTACCAAATGCCAAATTCCAAAGCCAGTGTGTGTGAAAGAAAGACCTTAGAAAAAGTCCA-	146				
QY	1952	ACTGTCATGGCCCATGTGTACAAAGATTGCCCTACAGATGAATCTGCAAGATGGGAGGAG	2011				
Db	147	--GGCAGGACATCGTCCACCAAGATTGCCCCAGCAGATGAATTGCAAGATGGGAGGCC	204				
QY	2012	CTCTGGAGG	2020				
Db	205	CTCTGGAGG	213				

RESULT 14
US-09-864-761-4964
; Sequence 4964, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6

```

1  / PRIOR FILING DATE: 2000-10-04
2  / PRIOR APPLICATION NUMBER: US 60/236,359
3  / PRIOR FILING DATE: 2000-09-27
4  / PRIOR APPLICATION NUMBER: PCT/US01/00666
5  / PRIOR FILING DATE: 2001-01-30
6  / PRIOR APPLICATION NUMBER: PCT/US01/00667
7  / PRIOR FILING DATE: 2001-01-30
8  / PRIOR APPLICATION NUMBER: PCT/US01/00664
9  / PRIOR FILING DATE: 2001-01-30
10 / PRIOR APPLICATION NUMBER: PCT/US01/00669
11 / PRIOR FILING DATE: 2001-01-30
12 / PRIOR APPLICATION NUMBER: PCT/US01/00665
13 / PRIOR FILING DATE: 2001-01-30
14 / PRIOR APPLICATION NUMBER: PCT/US01/00668
15 / PRIOR FILING DATE: 2001-01-30
16 / PRIOR APPLICATION NUMBER: PCT/US01/00663
17 / PRIOR FILING DATE: 2001-01-30
18 / PRIOR APPLICATION NUMBER: PCT/US01/00662
19 / PRIOR FILING DATE: 2001-01-30
20 / PRIOR APPLICATION NUMBER: PCT/US01/00661
21 / PRIOR FILING DATE: 2001-01-30
22 / PRIOR APPLICATION NUMBER: PCT/US01/00670
23 / PRIOR FILING DATE: 2001-01-30
24 / PRIOR APPLICATION NUMBER: US 60/234,687
25 / PRIOR FILING DATE: 2000-09-21
26 / PRIOR APPLICATION NUMBER: US 09/608,408
27 / PRIOR FILING DATE: 2000-06-30
28 / PRIOR APPLICATION NUMBER: US 09/774,203
29 / PRIOR FILING DATE: 2001-01-29
30 / NUMBER OF SEQ ID NOS: 49117
31 / SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
32 / SEQ ID NO 4964
33 / LENGTH: 364
34 / TYPE: DNA
35 / ORGANISM: Homo sapiens
36 / FEATURE:
37 / OTHER INFORMATION: MAP TO AP000359.1
38 / OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.9
39 / OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
40 / OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
41 / US-09-864-761-4964

```

	Query Match	2.0%;	Score 68;	DB 10;	Length 364;
	Best Local Similarity	62.2%;	Pred. No. 3.9e-06;		
	Matches 107;	Conservative	0;	Mismatches 65;	Indels 0; Gaps
QY	889	TCCAAGTTCACAGGTGGTGATTTGGCGCTGGCTTCACTACTTCCATCCTTCAGTAGTGA	948		
DB	193	TTCTAGTACAGGTTTGGAAATCTGCGTTGGTTATGTTACTTCTGTTCTTCAATACGAAA	252		
QY	949	CAGCATCATGCTCTGCACCTAGCGTTAGCGCATAAAGTCCCTTCGAAAGTCAGACTGTTTGG	1008		
DB	253	CAGCATTAACCTCTGTGCGGATGTGAGGCCANAACTGCTCGAATAGAACTGCTTATGA	312		
QY	1009	TTTTCATGTTCACTTTTATCATCAGACAGAAACATATAATTTTCAAGAACAA	1060		
DB	313	TTTTCATAAGAGAACATCTGCCAGGCCACAGGAGAAACATCCGACAGAA	364		

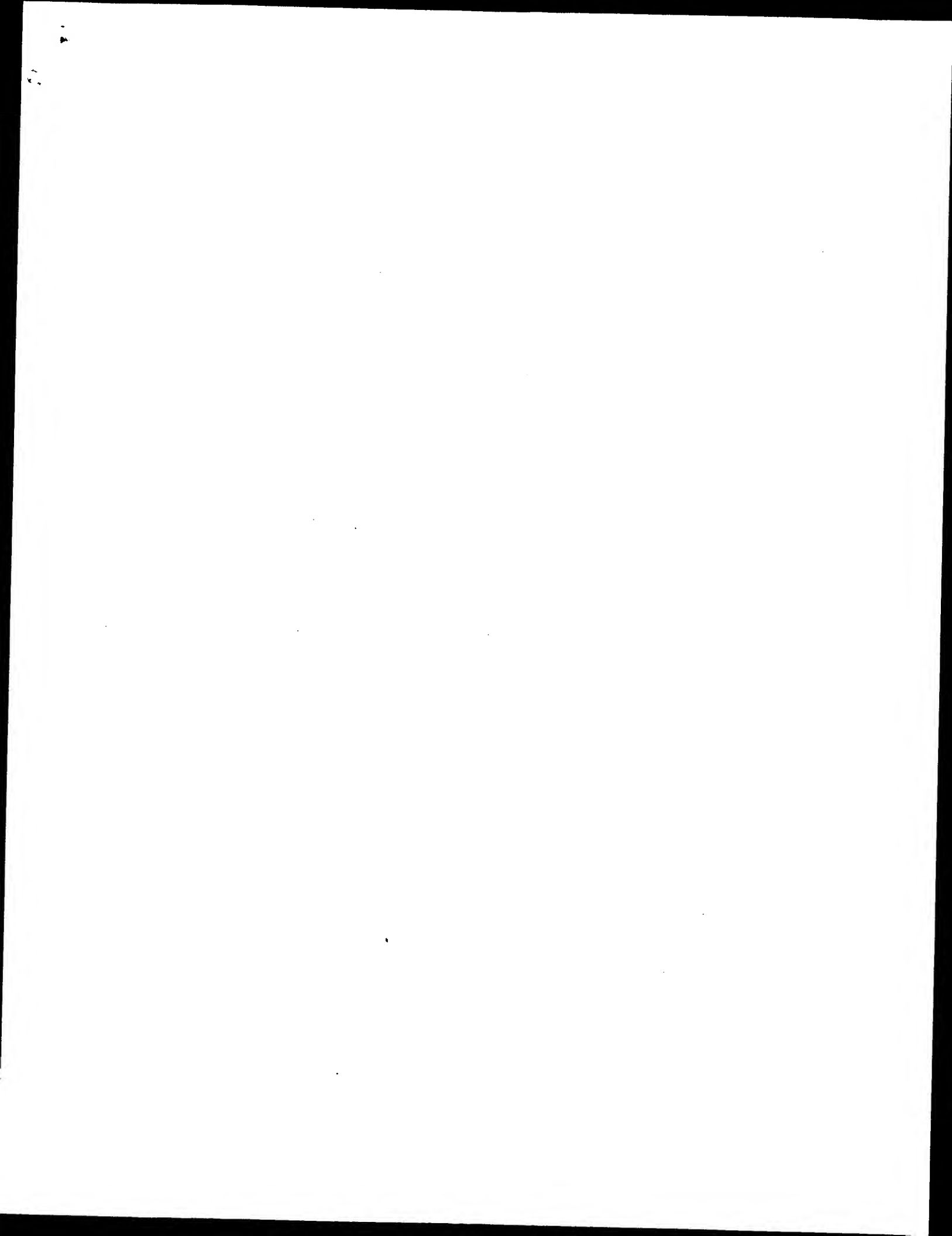
RESULT 15
US-09-960-352-3400
; Sequence 3400, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND COMPOSITIONS
; FILE REFERENCE: 16511.006/37-21(10298)
; CURRENT APPLICATION NUMBER: US/09/960,
; CURRENT FILING DATE: 2001-09-24

NUMBER OF SEQ ID NOS: 15112
 SEQ ID NO 3400
 LENGTH: 446
 TYPE: DNA
 ORGANISM: Bos taurus
 OTHER INFORMATION: Clone ID: 15-LIB3058-052-Q1-K1-D11
 US-09-960-352-3400

Query Match 1.9%; Score 67; DB 10; Length 446;
 Best Local Similarity 48.3%; Pred. No. 7.2e-06;
 Matches 187; Conservative 0; Mismatches 200; Indels 0; Gaps 0;

QY 2768 GCAGCGGCTTCTTCTTGAAGATGAGATTTTAACTTTTAACTTTT 2827
 DB 14 GCGGCTTCTTCTTGAAGATGAGATTTTAACTTTTAACTTTT 73
 QY 2828 TTTAACTGTTATCTTCTGAGGAACTTGGAGGGATAGAGATCTAGCATTTAT 2887
 DB 74 TTTTCTTCTTCTTCTGAGGAACTTGGAGGGATAGAGATCTAGCATTTAT 133
 QY 2888 TTTAGCATGCTATGACGGCTCTTATTTATATGTAATAATAGATTTTAT 2947
 DB 134 TTTTCTTCTTCTTCTGAGGAACTTGGAGGGATAGAGATCTAGCATTTAT 193
 QY 2948 TTTCTTCTTCTTCTGAGGAACTTGGAGGGATAGAGATCTAGCATTTAT 3007
 DB 194 TTTTCTTCTTCTTCTGAGGAACTTGGAGGGATAGAGATCTAGCATTTAT 253
 QY 3008 TGAATAGATCTGAGGAACTTGGAGGGATAGAGATCTAGCATTTAT 3067
 DB 254 TTTTCTTCTTCTTCTGAGGAACTTGGAGGGATAGAGATCTAGCATTTAT 313
 QY 3068 GATTACTTAAAGTCTTCTGAGGAACTTGGAGGGATAGAGATCTAGCATTTAT 3127
 DB 314 TTTTCTTCTTCTTCTGAGGAACTTGGAGGGATAGAGATCTAGCATTTAT 373
 QY 3128 TATTTTGAATCTGTTTGAATTT 3154
 DB 374 TTTTCTTCTTCTTCTGAGGAACTTGGAGGGATAGAGATCTAGCATTTAT 400

Search completed: January 9, 2003, 23:04:49
 Job time : 137 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2003, 18:54:20 ; Search time 3076 Seconds
(without alignments) 18280.471 Million cell updates/sec

Title: US-09-873-737a-5
Perfect score: 3472
Sequence: 1 gggcgcaattggccgcagcgt.....taataataattgctccccc 3472

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estlum:*
3: em_estlum:*
4: em_estlum:*
5: em_estlum:*
6: em_estlum:*
7: em_estlum:*
8: em_estlum:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estlum:*
16: em_estlum:*
17: gb_est1:*
18: em_gsa_hum:*
19: em_gsa_hum:*
20: em_gsa_hum:*
21: em_gsa_hum:*
22: em_gsa_hum:*
23: em_gsa_hum:*
24: em_gsa_hum:*
25: em_gsa_hum:*
26: em_gsa_hum:*
27: em_gsa_hum:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	736.8	21.2	818	BG773137 602721862
2	730.2	21.0	779	B1560710 603203560
3	689.2	19.9	1092	B1463482 603200456
4	672.2	19.4	882	B1459464 602695155
5	637.2	18.4	715	BG721550 603202635
6	633	18.2	666	B1463897 603202635

7	626.2	18.0	731	B1458664
8	621.4	17.9	640	AL705301
9	621	17.9	622	AL705072
10	580.6	16.7	718	BQ435882
11	574.6	16.5	874	BF666315
12	560.6	16.1	581	BG718168
13	523.2	15.1	539	B1463615
14	521	15.0	744	B1459142
15	515.2	14.8	577	BF080918
16	513.4	14.5	527	BF751055
17	502.4	14.5	1495	BF649841
18	492.8	14.2	605	AL704986
19	492.6	14.2	705	BF246708
20	471.2	13.6	533	BF080923
21	468.2	13.5	830	BF699368
22	452.4	13.0	463	AA639672
23	450.4	12.8	482	AM189518
24	443.4	12.7	472	AL705611
25	442.2	12.7	806	BG724366
26	440.4	12.7	462	AA969938
27	434.4	12.5	446	AL125224
28	415.8	12.0	454	AL044133
29	414.2	11.9	565	BG101107
30	412	11.6	417	AL041473
31	403.4	11.2	477	AA969660
32	387.8	11.1	420	BQ374910
33	387	11.1	477	BM339361
34	386	11.1	587	AA430311
35	373.4	10.8	935	BF247005
36	345.2	9.9	485	BG718140
37	339	9.8	971	BB614499
38	328.2	9.5	349	AA904973
39	326.8	9.4	872	B1560401
40	287.2	8.3	609	BQ481304
41	274.8	7.9	559	BM728018
42	274.4	7.9	300	AI538315
43	265.4	7.6	286	AI698821
44	263.8	7.5	512	BM680592
45	259.2	7.5	617	BQ400023

ALIGNMENTS

RESULT 1
BG773137 818 bp mRNA EST 15-MAY-2001
602721862P1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4838918 5',
LOCUS
DEFINITION
ACCESSION
BG773137
VERSION
BG773137.1 GI:14083790
KEYWORDS
SOURCE
ORGANISM
human.
REFERENCE
1 (bases 1 to 818)
NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
COMMENT
Contract: Robert Straube, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Tohiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL0773 row: k column: 15
High quality sequence stop: 789.
Location/Qualifiers

FEATURES

348 AGTCAAGAGACTCGAGATATCTGCTGATTTGAGAGTTATCGNTAGCAGAGAGAG 407
 Db 260 AGTCAAGAGACTCGAGATATCTGCTGATTTGAGAGTTATCGNTAGCAGAGAGAG 319
 Qy 408 GTCTGTGATGAGATTTTATGATCTTGGGNGAATPACAGAGAGAGAGAGAGAG 467
 Db 320 GTCTGTGATGAGATTTTATGATCTTGGGNGAATPACAGAGAGAGAGAGAGAG 379
 Qy 468 AAGATCAAAAACAGGTTCTTCAAGCATTTATAGTAAGGTTAAGCACTAACCTTTCCGGC 527
 Db 380 AAGATCAAAAACAGGTTCTTCAAGCATTTATAGTAAGGTTAAGCACTAACCTTTCCGGC 439
 Qy 528 TGACATCCCGTCCGAGGAGGCTTATATGATGATGATGATGATGATGATGATGATG 587
 Db 440 TGACATCCCGTCCGAGGAGGCTTATATGATGATGATGATGATGATGATGATGATG 499
 Qy 588 AAGCAGAGAGAGCTCGTTGATGATCTTCTTTTCAACAGAGAGATCTAATTTGAAAGTGC 647
 Db 500 AAGCAGAGAGAGCTCGTTGATGATCTTCTTTTCAACAGAGAGATCTAATTTGAAAGTGC 559
 Qy 648 ATGCTTTGATGAGATTTATTTATTTTCTTAAAGAGATGATGATGATGATGATG 707
 Db 560 ATGCTTTGATGAGATTTATTTATTTTCTTAAAGAGATGATGATGATGATGATG 619
 Qy 708 TTTTAAAG 767
 Db 620 TTTTAAAG 679
 Qy 768 TTCCAGCTTACATCAACATCTGTTTGGAGTTCTTAAATTTATTTTCAAGAGAGCTTTTGA 827
 Db 680 TTCCA-CTACATCAAGCACTGTTTGGAGTTCTTAAATTTATTTTCAAGAGAGCTTTTGA 738
 Qy 828 AAATCA-TGAATTTGCAACAATTTGAGCAAAATTTATTA 866
 Db 739 AAATCAATTTGCAACAATTTGAGCAAAATTTATTA 778

RESULT 3
 BI463482
 LOCUS 1092 bp mRNA linear EST 21-AUG-2001
 DEFINITION 603203560P1 NIH_MGC_97 Homo sapiens CDNA clone IMAGE:5269311 5',
 mRNA sequence.
 BI463482
 ACCESSION BI463482.1 GI:15254125
 VERSION EST.
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1092)
 NIH-MGC http://mgi.nci.nih.gov/
 COMMENT Unpublished (1999)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgsapb-r@mail.nih.gov
 Tissue Procurement: Miklos Palkeovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHRI), Shitaki
 Toshitaki and Piero Carninci (RIKEN)
 cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Inocyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: LLM11679 row: h column: 16
 High quality sequence stop: 668.
 Location/Qualifiers
 1..1092
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5269311"
 /clone_lib="NIH_MGC_97"
 /lab_host="DH10B"
 /note="Organ: testis; Vector: pBluescriptR (modified

Bluescript KS+; Site 1: BamHI; Site 2: SalI-XhoI (gcgcag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.2 kb and
 normalized to R0.5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NHRI/NHRI, National
 Institutes of Health). Note: this is a NIH-MGC Library."
 BASE COUNT 356 a 201 c 270 g 264 t 1 others
 ORIGIN

Query Match 19.9%; Score 689.2; DB 13; Length 1092;
 Best Local Similarity 83.5%; Pred. NO. 2.1e-136;
 Matches 903; Conservative 0; Mismatches 156; Indels 22; Gaps 10;

Qy 106 CAGGTGTGAGTGTGCAAGACCAAGACCTAGAGAGAGAGAGAGAGAGAGAGAG 165
 Db 16 CCGGAGTGTGAGTGTGCAAGACCAAGACCTAGAGAGAGAGAGAGAGAGAGAG 75
 Qy 166 AATGACTGTGAG 225
 Db 76 AAGACTGTGAG 135
 Qy 226 G-CTGGTGGGCTCCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 284
 Db 136 GACTGTGGGCTCCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 195
 Qy 285 CACCGAG 344
 Db 196 CACCGAG 255
 Qy 345 CCAAGTCAAG 404
 Db 256 CCAAGTCAAG 315
 Qy 405 GAGGTGTGAG 464
 Db 316 GAGGTGTGAG 375
 Qy 465 TTAAGATCAAAAACAGGTTCTTCAAGCATTTATAGTAAGTATGACATTTTCC 524
 Db 376 TTAAGATCAAAAACAGGTTCTTCAAGCATTTATAGTAAGTATGACATTTTCC 435
 Qy 525 GGTGATCATCCGTCCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 584
 Db 436 GGTGATCATCCGTCCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 495
 Qy 585 TGAAGCCAG 644
 Db 496 TGAAGCCAG 555
 Qy 645 GTGATGCTTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 704
 Db 556 GTGATGCTTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 615
 Qy 705 AAGTTTGTGAG 764
 Db 616 AAGTTTGTGAG 675
 Qy 765 AACTTCACTTATCAACCACTGTTTGTGAGAGAGAGAGAGAGAGAGAGAGAG 824
 Db 676 AACTTCACTTATCAACCACTGTTTGTGAGAGAGAGAGAGAGAGAGAGAGAG 732
 Qy 825 TGAAGATCAATTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 881
 Db 733 TGAAGATCAATTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 792
 Qy 882 TTAATTTTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 933
 Db 793 TTAATTTTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 852
 Qy 934 CTTGAGTGTGAG 989

Unpublished (1999)
Contact: Robert St

Unpublished (1999)
 Contact: Robert Strausberg, Ph. D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroaki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (LINTL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINTL at:
<http://image.llnl.gov>
 Plate: LLAM10743 row: m column: 14
 High quality sequence stop: 715.
cgabs-remail.nih.gov

location/Qualifiers
1. .715

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4827445"
/clone_1ib="NIH_MGC_97"
/lab_host="DH10B"
/notice="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI, Site_2: SalI-XhoI (gtcgagc
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTN-3',
size-selected for average insert size 2.2 kb and
normalized for ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
a 160 C 189 g 154 t

```

212 a 160 c 189 g 154 t

Query Match	18.4%;	Score 637.2;	DB 12;	length 715;
Best Local Similarity	98.1%;	Pred. No. 2.6e-125;		
Matches 686;	Conservative	0;	Mismatches 9;	Indels 4;
			Gaps	4.

QY	109	17	Db
	GTCGAGGCGCAAGCAACGAGACTAGGGCGAGGCGCGTCCAGAAATAGAAAAACAT	GCGTCGAGGCGCAAGCAACGAGACTAGGGCGAGGCGCGTCCAGAAATAGAAACAAAT	
	168	76	

QY	169	77	Db
GACTGGGAGAGCCCGAGCCAGAGCCAGAGGAGGGCCCCGGGTCAAGAGACAGCCAGCT	228	GACTGGGAGAGCCCGAGCCAGAGCCAGAGGAGGGCCCCGGGTCAAGAGACAGCCAGCT	136

QY	229	GGTGGGCTC	ACTGCACG	AGCAACCTG	TATATTTC	TAGCCTAGG	CCCTCAG	-CGGCAC	287
Db	137	GGTGGGCTC <td>ACTGCAGT <td>GCAACCTG <td>TATATTTC <td>AGCCTAGG <td>CCCTCAG <td>-CGGCAC</td> <td>196</td> </td></td></td></td></td>	ACTGCAGT <td>GCAACCTG <td>TATATTTC <td>AGCCTAGG <td>CCCTCAG <td>-CGGCAC</td> <td>196</td> </td></td></td></td>	GCAACCTG <td>TATATTTC <td>AGCCTAGG <td>CCCTCAG <td>-CGGCAC</td> <td>196</td> </td></td></td>	TATATTTC <td>AGCCTAGG <td>CCCTCAG <td>-CGGCAC</td> <td>196</td> </td></td>	AGCCTAGG <td>CCCTCAG <td>-CGGCAC</td> <td>196</td> </td>	CCCTCAG <td>-CGGCAC</td> <td>196</td>	-CGGCAC	196

QY	288	CAGCAGAGGGGAAATTATTTGGCCCTGAGACGCAGAGCAACAGCAGGAAACAGCCA	347
Db	197	CAGCAGAGGGGAAATTATTTGGCCCTGAGACGCAGAGCAACAGCAGGAAACAGCCA	256

Oy 348 AGTCAAGAAGACTCCAGATATCTGCTGGAGTTTCAGAGATTATCGNATAGCAGAGAAAGAG 407
 Db 257 AGTCAAGAAGACTCCAGATATCTGCTGGAGTTTCAGAGATTATCGTATAGCAGAGAGGAG 316

Oy	408	GTCTGCTGAGATTTTCATGATCTTGTGTGAATCAAGCGAACTTGACCATGTTA	467
Db	317	GTCTGCTGAGATTTTCATGATCTTGTGTGAATCAAGCGAACTTGACCATGTTA	376

QY 468 AAGATCAAAACAGGTCCTCAGGCATTATAGTAAGTTAAGCACTAACCATTTCCGC 527

Dh 377 AAGATCAAAACAGGTCCTCAGGCATTATAGTAAGTTAAGCACTAACCATTTCCGC 436

QY 528 TGACATCCCGTCCCACTGGGCTTATATAGATATACATTTGACATTAACCCACCATGATGG 587

Dh 437 TGACATCCCGTCCCACTGGGCTTATATAGATATACATTTGACATTAACCCACCATGATGG 496

QY 588 AAGCCGAGAGACTCCGCTTCACTCTTTTTCACACGAGATCTAATTGGAAAGTGTC 647

Dh 497 AAGCCGAGAGACTCCGCTTCACTCTTTTTCACACGAGATCTAATTGGAAAGTGTC 556

QY 648 ATGCTTTGATGGAACGATATTATTTTACCCTAAAAGACTACGCAAAA-GSTTACTGAA 706

Db 557 ATGCTTTGATGGAACGATATTATTATTACCTAAAAAGACTACAGCAAACGGTTACTGAA 616

Db 617 GTTATAGTAGACCCCGAATGAGAGATGTGAGATCACCGATCACTTTAACTG 676

677 AACCTCAGCTACATCACCACCTGTGTGCGGCTTATA 715
 Db

DEFINITION	603202635F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5268465 5', mRNA sequence.
ACCESSION	U1463897

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
Mammalia; Eutheria; Primates; Catarrhini, Homnidae; Homo.
1 (bases 1 to 666)
NIH-MGC <http://mcc.nci.nih.gov/>.

AUTHOR: Robert Strausberg, Ph. D.
 CONTACT: Robert Strausberg, Ph. D.
 EMAIL: rstraus@mail.nih.gov
 COMMENT: Unpublished (1999)
 JOURNAL: *Journal of Interpersonal Violence*
 YEAR: 2002

1986a, 1986b, 1986c, 1986d, 1986e, 1986f, 1986g, 1986h, 1986i, 1986j, 1986k, 1986l, 1986m, 1986n, 1986o, 1986p, 1986q, 1986r, 1986s, 1986t, 1986u, 1986v, 1986w, 1986x, 1986y, 1986z, 1987a, 1987b, 1987c, 1987d, 1987e, 1987f, 1987g, 1987h, 1987i, 1987j, 1987k, 1987l, 1987m, 1987n, 1987o, 1987p, 1987q, 1987r, 1987s, 1987t, 1987u, 1987v, 1987w, 1987x, 1987y, 1987z, 1988a, 1988b, 1988c, 1988d, 1988e, 1988f, 1988g, 1988h, 1988i, 1988j, 1988k, 1988l, 1988m, 1988n, 1988o, 1988p, 1988q, 1988r, 1988s, 1988t, 1988u, 1988v, 1988w, 1988x, 1988y, 1988z, 1989a, 1989b, 1989c, 1989d, 1989e, 1989f, 1989g, 1989h, 1989i, 1989j, 1989k, 1989l, 1989m, 1989n, 1989o, 1989p, 1989q, 1989r, 1989s, 1989t, 1989u, 1989v, 1989w, 1989x, 1989y, 1989z, 1990a, 1990b, 1990c, 1990d, 1990e, 1990f, 1990g, 1990h, 1990i, 1990j, 1990k, 1990l, 1990m, 1990n, 1990o, 1990p, 1990q, 1990r, 1990s, 1990t, 1990u, 1990v, 1990w, 1990x, 1990y, 1990z, 1991a, 1991b, 1991c, 1991d, 1991e, 1991f, 1991g, 1991h, 1991i, 1991j, 1991k, 1991l, 1991m, 1991n, 1991o, 1991p, 1991q, 1991r, 1991s, 1991t, 1991u, 1991v, 1991w, 1991x, 1991y, 1991z, 1992a, 1992b, 1992c, 1992d, 1992e, 1992f, 1992g, 1992h, 1992i, 1992j, 1992k, 1992l, 1992m, 1992n, 1992o, 1992p, 1992q, 1992r, 1992s, 1992t, 1992u, 1992v, 1992w, 1992x, 1992y, 1992z, 1993a, 1993b, 1993c, 1993d, 1993e, 1993f, 1993g, 1993h, 1993i, 1993j, 1993k, 1993l, 1993m, 1993n, 1993o, 1993p, 1993q, 1993r, 1993s, 1993t, 1993u, 1993v, 1993w, 1993x, 1993y, 1993z, 1994a, 1994b, 1994c, 1994d, 1994e, 1994f, 1994g, 1994h, 1994i, 1994j, 1994k, 1994l, 1994m, 1994n, 1994o, 1994p, 1994q, 1994r, 1994s, 1994t, 1994u, 1994v, 1994w, 1994x, 1994y, 1994z, 1995a, 1995b, 1995c, 1995d, 1995e, 1995f, 1995g, 1995h, 1995i, 1995j, 1995k, 1995l, 1995m, 1995n, 1995o, 1995p, 1995q, 1995r, 1995s, 1995t, 1995u, 1995v, 1995w, 1995x, 1995y, 1995z, 1996a, 1996b, 1996c, 1996d, 1996e, 1996f, 1996g, 1996h, 1996i, 1996j, 1996k, 1996l, 1996m, 1996n, 1996o, 1996p, 1996q, 1996r, 1996s, 1996t, 1996u, 1996v, 1996w, 1996x, 1996y, 1996z, 1997a, 1997b, 1997c, 1997d, 1997e, 1997f, 1997g, 1997h, 1997i, 1997j, 1997k, 1997l, 1997m, 1997n, 1997o, 1997p, 1997q, 1997r, 1997s, 1997t, 1997u, 1997v, 1997w, 1997x, 1997y, 1997z, 1998a, 1998b, 1998c, 1998d, 1998e, 1998f, 1998g, 1998h, 1998i, 1998j, 1998k, 1998l, 1998m, 1998n, 1998o, 1998p, 1998q, 1998r, 1998s, 1998t, 1998u, 1998v, 1998w, 1998x, 1998y, 1998z, 1999a, 1999b, 1999c, 1999d, 1999e, 1999f, 1999g, 1999h, 1999i, 1999j, 1999k, 1999l, 1999m, 1999n, 1999o, 1999p, 1999q, 1999r, 1999s, 1999t, 1999u, 1999v, 1999w, 1999x, 1999y, 1999z, 2000a, 2000b, 2000c, 2000d, 2000e, 2000f, 2000g, 2000h, 2000i, 2000j, 2000k, 2000l, 2000m, 2000n, 2000o, 2000p, 2000q, 2000r, 2000s, 2000t, 2000u, 2000v, 2000w, 2000x, 2000y, 2000z, 2001a, 2001b, 2001c, 2001d, 2001e, 2001f, 2001g, 2001h, 2001i, 2001j, 2001k, 2001l, 2001m, 2001n, 2001o, 2001p, 2001q, 2001r, 2001s, 2001t, 2001u, 2001v, 2001w, 2001x, 2001y, 2001z, 2002a, 2002b, 2002c, 2002d, 2002e, 2002f, 2002g, 2002h, 2002i, 2002j, 2002k, 2002l, 2002m, 2002n, 2002o, 2002p, 2002q, 2002r, 2002s, 2002t, 2002u, 2002v, 2002w, 2002x, 2002y, 2002z, 2003a, 2003b, 2003c, 2003d, 2003e, 2003f, 2003g, 2003h, 2003i, 2003j, 2003k, 2003l, 2003m, 2003n, 2003o, 2003p, 2003q, 2003r, 2003s, 2003t, 2003u, 2003v, 2003w, 2003x, 2003y, 2003z, 2004a, 2004b, 2004c, 2004d, 2004e, 2004f, 2004g, 2004h, 2004i, 2004j, 2004k, 2004l, 2004m, 2004n, 2004o, 2004p, 2004q, 2004r, 2004s, 2004t, 2004u, 2004v, 2004w, 2004x, 2004y, 2004z, 2005a, 2005b, 2005c, 2005d, 2005e, 2005f, 2005g, 2005h, 2005i, 2005j, 2005k, 2005l, 2005m, 2005n, 2005o, 2005p, 2005q, 2005r, 2005s, 2005t, 2005u, 2005v, 2005w, 2005x, 2005y, 2005z, 2006a, 2006b, 2006c, 2006d, 2006e, 2006f, 2006g, 2006h, 2006i, 2006j, 2006k, 2006l, 2006m, 2006n, 2006o, 2006p, 2006q, 2006r, 2006s, 2006t, 2006u, 2006v, 2006w, 2006x, 2006y, 2006z, 2007a, 2007b, 2007c, 2007d, 2007e, 2007f, 2007g, 2007h, 2007i, 2007j, 2007k, 2007l, 2007m, 2007n, 2007o, 2007p, 2007q, 2007r, 2007s, 2007t, 2007u, 2007v, 2007w, 2007x, 2007y, 2007z, 2008a, 2008b, 2008c, 2008d, 2008e, 2008f, 2008g, 2008h, 2008i, 2008j, 2008k, 2008l, 2008m,

For ordering information can be
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/INTL at:
<http://image.jnl.gov>

FEATURES	source
file: <code>data1071.com.c000</code>	
High quality sequence stop: 666.	
Location/Qualifiers	
1..666	

Location/Qualifiers
1. .666

```

/organism="Homo sapiens"
/db xref="taxon:9606"
/clone IMAGE:5268465"
/clone_1db="NIMH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgga
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.2 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIM_MGC library."

```

194 a	144 c	187 g	141 b
-------	-------	-------	-------

Query Match	18.2%	Score 633	DB 13	Length 666
Best Local Similarity	98.2%	Pred. No. 2.1e-124		
Matches 650; Conservative	0	Mismatches 11	Indels 1	Gaps 1

93 GCGGCGCGCGCGCGAGGTCTGAGGTGCAAGACCAAGGACAGGGGAGCGGATCCA 132
5 GGGGCTGTGGCCCTGGGGCTGAGGTCAGGACCAAGGACCAAGGACAGGGGAGCGGTCCA 64

02 AGAATAGAAAACATGACTGGGAGACCCGAGCCAGAGCAGAGCAGAGGAGGAGGATC 212
 153 AGAATAGAAAACATGACTGGGAGACCCGAGCCAGAGCAGAGCAGAGGAGGAGGATC 212
 Db 65 AGAATAGAAAACATGACTGGGAGACCCGAGCCAGAGCAGAGGAGGAGGATC 124

213 AGGAGCAGCCAGCGTGTGGGCTCCATCGCAGTCAGCAGCAACCTGGTATATACACCTA 212

Db 125 AGGAGACAGCGCAGCTGGTGGCTCCACTGCCAGTCAGCAACCTGGTTATATTACAGCCTA 184
 Qy 273 GGCCTCAGCCGCCACACAGAGGGGGAAATATTGGCCGTGGAGCGCAGAGAGGAACAG 332
 Db 185 GGCCTCAGCCGCCACACAGAGGGGGAAATATTGGCCGTGGAGCGCAGAGAGGAACAG 244
 Qy 333 CAGGAGGAACAGCCAAAGTCACAGAGCTCCAGATATCTGCTGGATTTTCAGAGTTATCGN 392
 Db 245 CAGGAGGAACAGCCAAAGTCACAGAGCTCCAGATATCTGCTGGATTTTCAGAGTTATCGT 304
 Qy 393 TAGCAGAGAGAGGAGTGGTGGTGGATTTTCAGATATCTGCTGGATTTTCAGAGTTATCGN 452
 Db 305 TAGCAGAGAGAGGAGTGGTGGTGGATTTTCAGATATCTGCTGGATTTTCAGAGTTATCGT 364
 Qy 453 ACCTAGAGCACTGTAAAGAAATCAAAACAGAGTCTTTCAGGCATATATAGTAAAGTAAAGCA 512
 Db 365 ACCTAGAGCACTGTAAAGAAATCAAAACAGAGTCTTTCAGGCATATATAGTAAAGTAAAGCA 424
 Qy 513 CTAACCACTTTCCGGCTGACATCCCGTCCCAAGTGGGCTTTATATCATGATTCACATTGACT 572
 Db 425 CTAACCACTTTCCGGCTGACATCCCGTCCCAAGTGGGCTTTATATCATGATTCACATTGACT 484
 Qy 573 ATAAACCACTGATGAA- GCCAGAGACTCCGTTTCAGCTCTTCTTTTCAACAGCAAGAT 631
 Db 485 ATAAACCACTGATGAAAGCGCCAGAGACTCCGTTTCAGCTCTTCTTTTCAACAGCAAGAT 544
 Qy 632 CTAATTCGAAAGTGTATGCTTTTATGAGAACCGATATATTTTACCTAAAGACTACAG 691
 Db 545 CTAATTCGAAAGTGTATGCTTTTATGAGAACCGATATATTTTACCTAAAGACTACAG 604
 Qy 692 CAAAGGTTACTGAAGTTTATGAGAACCGGATGAGAGGATGTAGGATAACGATC 751
 Db 605 CAAAGGTTACTGAAGTTTATGAGAACCGGATGAGAGGATGTAGGATAACGATC 664
 Qy 752 AC 753
 Db 665 AC 666

RESULT 7
 BI458664
 LOCUS 603199792F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5265838 5',
 DEFINITION mRNA sequence.
 ACCESSION BI458664
 VERSION BI458664.1 GI:15249320
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 731)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLNL1670 row: g column: 23
 High quality sequence stop: 671.
 Location/Qualifiers
 1. .731
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5265838"
 /clone_lib="NIH_MGC_97"

FEATURES

source
 RESULT 8
 AL705301
 LOCUS
 DEFINITION DKF2686G1235_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
 ACCESSION DKF2686G1235 5', mRNA sequence.
 VERSION AL705301
 GI:19688656

/lab_host="DH10B"

/note="Organ: testis; Vector: pBluescriptR (modified
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.2 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NHGRI/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 BASE COUNT 214 a 157 c 201 g 159 t
 ORIGIN

Query Match 18.0%; Score 626.2; DB 13; Length 731;
 Best Local Similarity 97.5%; Pred. No. 5.8e-123;
 Matches 667; Conservative 0; Mismatches 14; Indels 3; Gaps 3;

Qy 93 GGGGCGCCCGCGGAGGCTCAGGTGCAAGCAGCAGGACTAGGGGAGGCGGAGGCGGTCCA 152
 Db 18 GGGGCTGTTGGCTCGGGCTGAGGTGCAAGCAGCAGGACTAGGGGAGGCGGAGGCGGTCCA 77
 Qy 153 AGAAATAGAAAAACAATGACTGGGAGAGCCCGAGCAGAGCCAGAGGAGGCGCCCGGTC 212
 Db 78 AGAAATAGAAAAACAATGACTGGGAGAGCCCGAGCAGAGCCAGAGGAGGCGCCCGGTC 137
 Qy 213 AGGAGACAGCGAGCTGGTGGCTCCACTGCCAGTCAGCAACCTGGTTATATTACAGCCTA 272
 Db 138 AGGAGACAGCGAGCTGGTGGCTCCACTGCCAGTCAGCAACCTGGTTATATTACAGCCTA 197
 Qy 273 GGCCTCAGCCGCCACACAGCAGAGGGGAAATATTGGCCGTGGAGCGCAGAGAGGAACAG 332
 Db 198 GGCCTCAGCCGCCACACAGCAGAGGGGAAATATTGGCCGTGGAGCGCAGAGAGGAACAG 257
 Qy 333 CAGGAGGAACAGCCAAAGTCACAGGACTCCAGATATCTGCTGGATTTTCAGAGTTATCGT 392
 Db 258 CAGGAGGAACAGCCAAAGTCACAGGACTCCAGATATCTGCTGGATTTTCAGAGTTATCGT 317
 Qy 393 TAGCAGAGAGAGGAGTGGTGGTGGATTTTCATGATCTTGGTGTGAATACAAGGCAGA 452
 Db 318 TAGCAGAGAGAGGAGTGGTGGTGGATTTTCATGATCTTGGTGTGAATACAAGGCAGA 377
 Qy 453 ACCTAGACCATGTTAA-AGAATCAAAACAGGTTCTTCAGGCATATATAGTAAAGTTAAGC 511
 Db 378 ACCTAGACCATGTTAAAGATCAAAACAGGTTCTTCAGGCATATATAGTAAAGTTAAGC 437
 Qy 512 ACTAACCATTTCCGGCTGACATCCCGTCCCGAGTGGGCTTTATATCATGATTCACATTGAC 571
 Db 438 ACTAACCATTTCCGGCTGACATCCCGTCCCGAGTGGGCTTTATATCATGATTCACATTGAC 497
 Qy 572 TATAACCCACTGATGGAAGCCAGAGACTCCGTTACGCTCTCTTTTCAACACGAAGAT 631
 Db 498 TATAACCCACTGATGGAAGCCAGAGACTCCGTTACGCTCTCTTTTCAACACGAAGAT 557
 Qy 632 CTAATTCGAAAGTGTCACTGCTTTTGTAGGAACAGATATATTTTACCTAAAGACTACAG- 690
 Db 558 CTAATTCGAAAGTGTCACTGCTTTTGTAGGAACAGATATATTTTACCTAAAGACTACAG 617
 Qy 691 GCATAAGGTTACT-GAAGTTTGTAGTAAGACCCGGAATGGAGAGGATGTGAGGATAACGA 749
 Db 618 GCATAAGGTTACTGGAAGTTTGTAGTAAGACCCGGAATGGAGAGGATGTGAGGATAACGA 677
 Qy 750 TCACCTTTAAACAATGACTTCCAC 773
 Db 678 TCACCTTTAAACAATGACTTCCAC 701

AL705301 640 bp mRNA linear EST 22-MAR-2002
 DKF2686G1235_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
 DKF2686G1235 5', mRNA sequence.
 AL705301
 AL705301.1 GI:19688656

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 640)
Ostenwaelder, B., Obermaier, B., Mewes, W., Mewes, H. W., Weill, B. and Wiemann, S.
TITLE EST (Ostenwaelder, B., Obermaier, B., Mewes, H. W., Weill, B. and Wiemann, S.)
JOURNAL Unpublished (2001)
COMMENT Contact: Ostenwaelder B
MIPS
Am Kioferplatz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de
sequenced by Medigenomix (Martinsried/Germany) within the CDNA sequencing consortium of the German Genome Project. No si sequence available.
This clone (DKFZ68661235) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..640
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZ68661235"
/clone_lib="686 (synonym: hlcc3)"
/tissue_type="human skeletal muscle"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pTriplEx2; Site_1: SfiI; Site_2: SfiI; CDNA-collection"
BASE COUNT 190 a 139 c 174 g 137 t
ORIGIN
Query Match 17.9%; Score 621.4; DB 9; Length 640;
Best Local Similarity 98.9%; Pred. No. 6,2e-122;
Matches 625; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 108 GGCTGTAGGTGCAAGACCAAGACTGAGGCGAGGCGGCTCCAGAAATGAAAAACA 167
DB 9 GGGCTGAGGTGCAAGACCAAGACTGAGGCGAGGCGGCTCCAGAAATGAAAAACA 68
QY 168 TGACTGGAGAGCCCGACCAAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 227
DB 69 TGACTGGAGAGCCCGACCAAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 128
QY 228 TGGTGGGCTCCATGCGCAAGTCAAGTCTGTTATATTCAGCTTACGCGCCAC 287
DB 129 TGGTGGGCTCCATGCGCAAGTCAAGTCTGTTATATTCAGCTTACGCGCCAC 188
QY 288 CAGCAGAGGGGAGATTTATTTGGCCGTGACGGCAGAGAGAACAGAGAACAGCA 347
DB 189 CAGCAGAGGGGAGATTTATTTGGCCGTGACGGCAGAGAGAACAGAGAACAGCA 248
QY 348 AGTCAACAAGGACTCCAGATATCTGCTGAGTTTCAAGAGTTATCGTAAAGAGAGAG 407
DB 249 AGTCAACAAGGACTCCAGATATCTGCTGAGTTTCAAGAGTTATCGTAAAGAGAGAG 308
QY 408 GTGCTGTAGAGATTTTTCATGATCTTGTGTGATAACAAGCAGAACTTACATGTTA 467
DB 309 GTGCTGTAGAGATTTTTCATGATCTTGTGTGATAACAAGCAGAACTTACATGTTA 368
QY 468 AAGATATTAACAGAGTTCTTACGAGTATTAAGAGTTAAGCACTTACATTTCCGCG 527
DB 369 AAGATATTAACAGAGTTCTTACGAGTATTAAGAGTTAAGCACTTACATTTCCGCG 428
QY 528 TGACATCCGCTCCAGTGGGCTTATATCAGTATCAGATTCATGACTTAAACCACTGATG 587
DB 429 TGACATCCGCTCCAGTGGGCTTATATCAGTATCAGATTCATGACTTAAACCACTGATG 488

QY 588 AGCCAGAAAGACTCCGTTAGACTCTTCTTTTCAACAGAAATTAATGGAAGTGTG 647
DB 489 AAGCCAGAAAGACTCCGTTAGACTCTTCTTTTCAACAGAAATTAATGGAAGTGTG 548
QY 648 ATGCTTTTGAATGAGACATTTATTTTACCTTAAAGACTTACAGCAAAAGCTTACGAG 707
DB 549 ATGCTTTTGAATGAGACATTTATTTTACCTTAAAGACTTACAGCAAAAGCTTACGAG 608
QY 708 TTTTGTAGTAAGCCGGAATGAGAGATGTG 739
DB 609 TTTTGTAGTAAGCCGGAATGAGAGATGTG 640
RESULT 9
AL705072 622 bp mRNA linear EST 22-MAR-2002
LOCUS DKFZ686E0634.F1.686 (synonym: hlcc3) Homo sapiens CDNA clone
DEFINITION DKFZ686E0634.5', mRNA sequence.
ACCESSION AL705072
VERSION AL705072.1 GI:19688427
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 622)
Ostenwaelder, B., Obermaier, B., Mewes, W., Mewes, H. W., Weill, B. and Wiemann, S.
TITLE EST (Ostenwaelder, B., Obermaier, B., Mewes, W., Weill, B. and Wiemann, S.)
JOURNAL Unpublished (2001)
COMMENT Contact: Ostenwaelder B
MIPS
Am Kioferplatz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de
sequenced by Medigenomix (Martinsried/Germany) within the CDNA sequencing consortium of the German Genome Project. No si sequence available.
This clone (DKFZ686E0634) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..622
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZ686E0634"
/clone_lib="686 (synonym: hlcc3)"
/tissue_type="human skeletal muscle"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pTriplEx2; Site_1: SfiI; Site_2: SfiI; CDNA-collection"
BASE COUNT 187 a 134 c 168 g 133 t
ORIGIN
Query Match 17.9%; Score 621; DB 9; Length 622;
Best Local Similarity 99.8%; Pred. No. 7,6e-122;
Matches 621; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 119 CAAGACCAAGACTGAGGCGGAGGCGGCTCCAGAAATGAAAAATGACTGGAGAGA 178
DB 1 CAAGACCAAGACTGAGGCGGAGGCGGCTCCAGAAATGAAAAATGACTGGAGAGA 60
QY 179 GCCGAGGCAAGGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 238
DB 61 GCCGAGGCAAGGCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
QY 239 ACTGCAGTCAAGCACTGTTATATTCAGCTTATAGGCTTACGCGCCACCAAGAGAGG 298
DB 121 ACTGCAGTCAAGCACTGTTATATTCAGCTTATAGGCTTACGCGCCACCAAGAGAGG 180

Fri Jan 10 09:31:38 2003

constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 153 a 127 c 154 g 105 t

ORIGIN

Query Match 15.1%; Score 523.2; DB 13; Length 539;
Best Local Similarity 99.2%; Pred. No. 5.2e-101;
Matches 525; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 106 CAGGTCTCAGGTGCAAGGACCCAGGACTAGGGCGAGGCGGCGGCTCAAGAAATAGAAAC 165
Db 11 CGGGTTTCTGAGGTGCAAGGACCCAGGACTAGGGCGAGGCGGCTCAAGAAATAGAAAC 70
QY 166 AATGCTGGGAGGCGGCGGAGCCAGGCGGAGGAGGCGGCGGCTCAAGAAATAGAAAC 225
Db 71 AATGCTGGGAGGCGGCGGAGCCAGGCGGAGGAGGCGGCGGCTCAAGAAATAGAAAC 130
QY 226 GCTGGTGGGCTCCACTGCGAGTCAAGCACTGCTGTTATATTCAGCTTAGGCTCAGCGCC 285
Db 131 GCTGGTGGGCTCCACTGCGAGTCAAGCACTGCTGTTATATTCAGCTTAGGCTCAGCGCC 190
QY 286 ACCAGCAGAGGGGGAAATTTTGGCCGTGGAGCGGAGGAGGAGGAGGAGGAGGAGG 345
Db 191 ACCAGCAGAGGGGGAAATTTTGGCCGTGGAGCGGAGGAGGAGGAGGAGGAGGAGG 250
QY 346 CAAGTCAGAGGACTCCAGATATCTGCTGGATTTTCAGGAGTTATCGTTCAGGAGAGG 405
Db 251 CAAGTCAGAGGACTCCAGATATCTGCTGGATTTTCAGGAGTTATCGTTCAGGAGAGG 310
QY 406 AGTCTGCTAGAGATTTTCATGATCTTGGTGTGAATCAAGGAGGAGGAGGAGGAGGAGG 465
Db 311 AGTCTGCTAGAGATTTTCATGATCTTGGTGTGAATCAAGGAGGAGGAGGAGGAGG 370
QY 466 TAAAGAAATCAAAACAGGTTCTTCAGGCAATATAGTAGGTTTACGACTAACCACTTCCG 525
Db 371 TAAAGAAATCAAAACAGGTTCTTCAGGCAATATAGTAGGTTTACGACTAACCACTTCCG 430
QY 526 GCTGACATCCGTCCTCCAGTGGGCTTTATATCAGTATCAGTATCAGTATCAGTATCAGTAT 585
Db 431 GCTGACATCCGTCCTCCAGTGGGCTTTATATCAGTATCAGTATCAGTATCAGTATCAGTAT 490
QY 586 GGAAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 634
Db 491 GGAAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 539

RESULT 14
BI459142 744 bp mRNA linear EST 21-AUG-2001
LOCUS 603199321F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5265484 5',
DEFINITION mRNA sequence.
ACCESSION BI459142
VERSION BI459142.1 GI:15249798
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11686 row: k column: 18
High quality sequence stop: 539.
Location/Qualifiers
1..539
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5272073"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified pluescript KS+); Site: 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library

FEATURES
source

125 GAGACAGCGCAGCTGGTGGGCTCCACTGCCAGTCAGCAACCTGGTTATATTCAGCCTAGG 184
QY 275 CCTCAGCCGCCACACAGCAGAGGGGGAATTTATTTGGCCGTGGACCGGCGAGAGGAGGAGCA 334
Db 185 CCTCAGCCGCCACACAGCAGAGGGGGAATTTATTTGGCCGTGGACCGGCGAGAGGAGGAGCA 244
QY 335 GGAGGAAACAGGCAAGGTCACAGGACTCCAGATATCTGCTGGATTTTCAGGAGTTATCGTTA 394
Db 245 GGAGGAAACAGGCAAGGTCACAGGACTCCAGATATCTGCTGGATTTTCAGGAGTTATCGTTA 304
QY 395 GCAGAGAGGAGGCTCGTCTGAGAGATTTTCATGATCTTGGTGTGAATPACAGGAGCAAGC 454
Db 305 GCAGAGAGGAGGCTCGTCTGAGAGATTTTCATGATCTTGGTGTGAATPACAGGAGCAAGC 364
QY 455 CTAGACCATGTTAAAGAAATCAAAACAGGTTCTTCAGGCAATATAGTAGGTTAAGCACT 514
Db 365 CTAGACCATGTTAAAGAAATCAAAACAGGTTCTTCAGGCAATATAGTAGGTTAAGCACT 424
QY 515 AACCATTTCCGGCTGACATCCCGTCCCGAGTGGGCTTATATCAGTATCAGATTCAGTAT 574
Db 425 AACCATTTCCGGCTGACATCCCGTCCCGAGTGGGCTTATATCAGTATCAGATTCAGTAT 484
QY 575 AACCCACTGATGAAGCCAGAGACTCCGTTTCAGCTCTTCTTTTCAACACGAGATCTA 634
Db 485 AACCCACTGATGAAGCCAGAGACTCCGTTTCAGCTCTTCTTTTCAACACGAGATCTA 544
QY 635 ATTGGAAGTGTGATGTTTGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 670
Db 545 ATTGGAAGTGTGATGTTTGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 580

RESULT 13
BI463615 539 bp mRNA linear EST 21-AUG-2001
LOCUS 603206270F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5272073 5',
DEFINITION mRNA sequence.
ACCESSION BI463615
VERSION BI463615.1 GI:15254271
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11686 row: k column: 18
High quality sequence stop: 539.
Location/Qualifiers
1..539
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5272073"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified pluescript KS+); Site: 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC http://mgi.nci.nih.gov/
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshitsugu and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11686 row: k column: 18
High quality sequence stop: 539.
Location/Qualifiers
1..539
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5272073"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified pluescript KS+); Site: 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library

High quality sequence stop: 667.

FEATURES
Source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5265484"
/clone_1ib="NIH_MGC_97"
/lab_host="DH10B"

/note="Organ: testis; Vector: pBluescript (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag site)-dtr primed using primer 5'-TTTTTTTTTTTTTTTAA-3', size-selected for average insert size 2.2 kb and normalized to 10⁵. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC library."

BASE COUNT 226 a 155 c 207 g 156 t

Query Match 15.0%; Score 521; DB 13; Length 744;
Best Local Similarity 92.6%; Pred. No. 1.5e-100;
Matches 624; Conservative 0; Mismatches 41; Indels 9; Gaps 7;

103 GGGGAGGTCTGAGGTGAGAGACGAGACCTAGGCGGACCGTCCAGAAATAGAA 162
29 GCGTCGGGCTGAGGTGAGAGACGAGACCTAGGCGGACCGTCCAGAAATAGAA 88
163 AACATGACTGGAGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 222
89 AACATGACTGGAGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 148
223 GCGAGTGTGGGCTGAGGTGAGAGACGAGACCTAGGCGGACCGTCCAGAAATAGAA 282
149 GCGAGTGTGGGCTGAGGTGAGAGACGAGACCTAGGCGGACCGTCCAGAAATAGAA 207
283 GCCACGAG 341
208 GCCACGAG 267
342 CAGCAG 401
268 CAGCAG 327
402 GAGGAGGTGTGAG 461
328 GAGGAGGTGTGAG 387
462 ATGTTAAAGATCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 521
388 ATGTTAAAGATCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 447
522 TCCGCTGACATCCGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 579
448 TCCGCTGACATCCGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 507
580 ACTGATGAG 637
508 CGTGAATGAG 567
638 GGAAGGTGATGATCTTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 697
568 GGAAGGTGATGATCTTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 626
698 GTTACTGAAGTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 756
627 GTTACTGAAGTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 685
757 AACCAATGAATTC 770
686 AACCAATGAATTC 699

RESULT 15

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BF080918 577 bp mRNA linear EST 18-OCT-2000
233684 MABC 2P1G Sus scrofa cDNA 5', mRNA sequence.
BF080918
BF080918.1 GI:10874739
EST.
pig.
Sus scrofa
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Familiarius; S.C.; Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casae, E.,
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.,
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTCCTGTCAGCAGC
Plate: 53 row: N column: 21
Seq primer: ATTTAGGTGACACTATG.
Location/Qualifiers
1..577
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_1ib="MARC 2P1G"
/tissue type="pooled"
/lab host="DH10B"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

FEATURES

Source

BASE COUNT

198 a 98 c 132 g 149 t

Query Match 14.8%; Score 515.2; DB 12; Length 577;
Best Local Similarity 93.4%; Pred. No. 2.6e-99;
Matches 538; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

1556 CAGGTGGAAGAAACATTTGATCAATCCACATTTGACAGATTGGTCCAAAGAAACAG 1615
2 CAGGTGGAAGAAACATTTGATCAATCCACATTTGACAGATTGGTCCAAAGAAACAG 61
1616 GGTGACCATTAATTAAGTTAAAGCACTGATGATGCTGTGATCTTAAGCGGAG 1675
62 GGTGACCATTAATTAAGTTAAAGCACTGATGATGCTGTGATCTTAAGCGGAG 121
1676 AATTATGAGCAGCAATCATTTGATCAATCCACATTTGACAGATTGGTCCAAAGAAACAG 1735
122 AATTATGAGCAGCAATCATTTGATCAATCCACATTTGACAGATTGGTCCAAAGAAACAG 181
1736 AATCAATGAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1795
182 AATCAATGAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 241
1796 GTCTTACAGCAAAAGGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1855
242 GTCTTACAGCAAAAGGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 301
1856 CGAAGAGCAAAATAGCATCTATTAATAAATACCTGTGTGACAGATTGGCCTTACCAAGT 1915
302 CGAAGAGCAAAATAGCATCTATTAATAAATACCTGTGTGACAGATTGGCCTTACCAAGT 361

us-09-873-737a-5.rst

Fri Jan 10 09:31:38 2003

```

QY 1916 CAGTGTGTGGTGGCCCGAACCTTAGGCAACAGCAAACTGTCTATGGCCATTGCTACAAAG 1975
Db      |||||
QY 362  CAGTGTGTGTGGTGGCCCGAACCTTAGGCAACAGCAAACTGTCTATGGCCATTGCGACAAAG 421
Db      |||||
QY 1976 ATTGCCCTACAGATGAACCTGCAAGATGGGAGGAGAGCTCTGGAGGGTGGACATCCCCCTG 2035
Db      |||||
QY 422  ATCGCCCTGCAGATGNACTGCAAGATGGGAGGAGAGCTTTGGAGGGTTGATATGCTCTA 481
Db      |||||
QY 2036 AAGCTCGTCATGATCGTTGGCATTCGATTGTTACCATGACATGACAGCTGGGCGGAGGTCA 2095
Db      |||||
QY 482  AAGCTAGCGATGATAGTTGGCATTTGATTCTTATCATGATACCACAGCTGGACGGAGGTCA 541
Db      |||||
QY 2096 ATCGCAGGATTTGTTGCCAGCATCAATGAAGGATG 2131
Db      |||||
QY 542  ATTGCAGGATTTGTTGCAAGCATCAATGAAGGATG 577
Db      |||||

```

Search completed: January 9, 2003, 21:34:36
Job time : 3089 secs